The application of statistical methods to problems relating to the chemical industry

Gent, M.

THE APPLICATION OF STATISTICAL METHODS TO PROBLEMS RELATING TO THE CHEMICAL INDUSTRY

by

M. GENT

AUGUST, 1960

SUMMARY

Most of the work recorded in this thesis was carried out during a year of postgraduate study at the University of Durham. There are three main sections in the thesis. One comprises an account of some field work carried out in a local industry, another is a review of some recently developed statistical methods which were applied to the practical problems, and the third is a development of some theory on missing values.

A study of a kiln process is described in which the effect of four factors on the quality of the final product were examined. An experiment on a related problem, on a laboratory scale, is described. In both of these investigations it was intended to determine the optimum working combination of a number of factors, and the methods of Box and Wilson were to be applied. The difficulties of adhering strictly to these methods in practice in these instances are described and how useful are the applications of their general principles. An appraisal of the work of Box and Wilson, together with a review of other related papers prefaces the practical results

Other kiln records showing the effects of several factors on the amount of dust lost were studied. The important factors were isolated and a relationship between them and the dust loss was estimated.

Some of the observations obtained in one of the kiln experiments were of doubtful validity because of changed operating conditions, and as a result some thought was given to the problem of estimating missing values in Factorial experiments. A method of estimating a single missing result in a balanced complete block design is described which is believed to be original. This technique enables a missing observation in a Factorial experiment to be estimated much more quickly than the current iterative method, and in addition it affords a new general method of deriving known formulae for other designs. A review of other methods is also given.

ACKNOWLEDGEMENTS.

The work presented in this thesis was carried out under the supervision of Dr. E.S. Page to whom I am very grateful for encouragement and advice. I am also indebted to the Management of the Steetley Magnesite Company for allowing me to help in the design and analysis of some of their research and development work, and to Dr. W.C. Gilpin and Mr. N. Heasman, who directed these experiments, for their co-operation and interest.

TABLE OF CONTENTS

CHAPTER 1.

THE EXPERIMENTAL DETERMINATION OF OPTIMUM CONDITIONS

1.1 INTRODUCTION

In 1951 Box and Wilson (1) read a paper before the Royal Statistical Society on the experimental determination of optimum conditions, with particular reference to chemical processes. Although the methods had been developed from the authors' experience in the chemical industry they have a more general application and the paper is a valuable contribution to industrial experimentation. It is intended to review this and other related papers in this chapter.

The work of Box and Wilson is about the design and analysis of experiments intended to explore functional relationships between a dependent variable, η , and several independent variables x_1, x_2, $\ldots x_k$. In the chemical industry the response to be optimised may be the yield or purity or cost and the corresponding variables may be temperature, concentration and time of reaction. It is assumed that the independent variables, or factors, can be varied on a continuous scale and that they may be controlled at any pre-assigned value. The response variable, η , is supposed dependent on these other variables according to an unknown function $\eta = \phi\ (x_1,x_2, \ldots x_k)$, which defines the Response Surface. For any combination of the x variables the observed response, y, will vary in repeated trials and have mean η and variance σ^2. The object of course, is to draw inferences about $\phi\ (x_1,x_2, \ldots x_k)$ from the observed response pattern and ultimately to find that combination of the factors which optimises the response within a given region in the k-factor space, defined by the practical limitations to changes in the factors.

One way of finding optimum conditions is to explore the whole experimental region but this would generally require a prohibitive number of experiments. Such a procedure is clearly inefficient as it would give a representation of the whole region when it is only in the neighbourhood of the optimum that an adequate representation is required. In practice the experimental error is often small, relative to the possible gains, and in addition the experiments can often be performed sequentially. A technique is therefore needed which will exploit these advantages. Box and Wilson proposed that as a first step a small group of experiments should be carried out at some base point corresponding to the best prior estimate of the optimum conditions. If in this sub-region the response surface can be represented locally by a hyperplane then the maximum increase in response can be obtained by following the calculated path of steepest ascent. This corresponds to altering each factor, x_i, in proportion to its estimated first derivative, b_i, from the base point at the centre of the design. A provisional optimum will be determined at some point along this path and this can be taken as the base point for another small group of experiments, from which a new path of steepest ascent is calculated. This procedure is continued until a region is reached in which second order effects become important. Here the response surface is locally nearly flat and the region is said to be near-stationary, however an exception is when a ridge of rising response is crossed. In many applications the initial sub-region may not be adequately represented by a hyperplane in which case a near stationary region will have been found without the use of the steepest ascent procedure.

The above technique has the disadvantage that it may find only a local maximum and as a result may miss a higher maximum. Where fuller exploration is impracticable this risk has to be accepted, but it is thought that surfaces with more than one peak are relatively rare.

Suitable chosen supplementary experiments will then be performed in the neighbourhood of the highest response obtained at this stage, which will allow an equation of second degree to be fitted. Analysis of this equation will indicate whether there is a local maximum or some ridge system. If an absolute maximum is indicated its coordinates can be determined and checked by carrying out a few confirmatory trials, but if there is no absolute maximum then the subsequent experiments will depend on the shape of the surface; if it is found that the surface cannot be adequately represented by an equation of second degree then it will be necessary to fit higher order terms. In practice however the quadratic response surface is often adequate.

Most of the papers on this work have been about the experimental designs used to estimate the functional relationship. This problem was dealt with only in part by Box and Wilson in their original paper who had until that time, relied mainly on conventional designs such as the factorial type, although in addition they had created a new "composite" design. Since then designs have been developed which are very efficient for deriving equations of up to order three. These designs were constructed to satisfy as many as possible of the following requirements:-

(a) The design should allow the graduating polynomial to represent the truefunction as accurately as possible within a specified region.

(b) It should allow a check to be made on the adequacy of the model.

(c) It should form a nucleus which can be augmented to satisfactory designs of higher order in case the model is inadequate.

(d) It should lend itself to "blocking".

(e) It should not contain an excessively large number of experimental points.

1.2. THE EXAMINATION OF RESPONSE SURFACES

DERIVATION OF THE EQUATION OF THE RESPONSE SURFACE.

The form of the response function $\phi(x_1, x_2, \ldots x_k)$ will be unknown but it is assumed that it can be represented by a Taylor Series of some order. The response surface is therefore represented by a regression equation of the form

$$\eta = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k}\sum_{j=1}^{k} \beta_{ij} x_i x_j + \cdots \qquad \text{------------} \quad (1)$$

If responses are observed at a suitably placed set of N points in the k-factor space estimates b_0, b_1, ..., of β_0, β_1, ... can be obtained by fitting the regression equations. This set of points constitutes the experimental design and defines the (N x k) design matrix. The observed responses at these points can be denoted by the N elements of a vector Y, and $\eta = E(Y)$ is the corresponding vector of expected values. If the equation of the response surface contains L terms and X is the (N x L) matrix of independent variables then

$$\eta = X\beta, \qquad \text{-------(2)}$$

where β is the vector of unknown constants.

The surface is fitted to the observed results by the method of least squares and if the errors in the responses are uncorrelated and have variance σ^2 and if in addition X has rank L, it is known that

(i) Unbiased estimates of β_i which are linear in the observations and have minimum variance are given by the elements of the vector $B = TY$, where the transforming matrix T is $(X'X)^{-1}X'$.

(ii) The variances and covariances of these estimates are the elements of the matrix $C^{-1}\sigma^2$, where $C^{-1} = (X'X)^{-1}$ and is called the precision matrix.

(iii) An unbiased estimate of $(N-L)\sigma^2$ is given by the residual sum of squares $(N-L)s^2 = (Y-XB)' (Y-XB) = Y'Y - Y'XB$.

METHOD OF STEEPEST ASCENT

In many practical situations the experimental error is small compared with the expected increase in response and in such circumstances it is often possible to exploit this by initially representing the surface in a sub-region of the factor space by an equation of first degree, namely

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots\cdots + \beta_k x_k. \qquad \text{----- (3)}$$

This of course represents a hyperplane and can be expected to provide a satisfactory approximation to the true surface only in regions of limited extent and removed from stationary points.

If a series of experiments is carried out in a region in which this approximation is adequate the pattern of the observed responses can be used to indicate the best direction in which to move for higher responses. The method employed is that of steepest ascent, in which the factors are varied in proportion to the estimated regression coefficients. Geometrically, this means the path followed is in the direction at right angles to contours of equal response which are assumed to be locally parallel and equidistant. Further experiments along this path can be expected to produce higher and higher responses until eventually an apparent stationary point will be reached. If at this point first order coefficients are still dominant further experiments and subsequent application of the method of steepest ascent will lead to still higher responses. This procedure is continued until first order effects can no longer be regarded as large compared with effects of higher order. At this stage a near-stationary region will have been reached.

If after the initial series of experiments, certain of the factors show only small effects, it may be that the factor has no real effect at all on the response or it may be due to either the mean level for the factor being chosen near a conditional maximum or the change in level for the factor being too small. To safeguard against a wrong decision the average level of the factor should be changed away from the calculated path of steepest ascent and an increase made in the difference between the levels of the factor. If the factor is still found to be without effect it can then be discarded, while if a real effect is found then it would appear that the earlier absence of effect had been due to one of the two latter reasons given above.

The above technique will not in itself locate a maximum but it will lead to a near stationary region. This region cannot contain a minimum but it need not necessarily contain a true maximum. Since only simple equations of first degree are fitted in the method of steepest ascent the number of experiments required is relatively small and it is thus possible to make rapid progress very economically. If the original sub-region cannot be adequately represented by a plane the above technique cannot be applied and it will be necessary to carry out further experiments in that region in order to fit an equation of higher degree.

* See facing page.

EXPLORATION OF THE SURFACE IN A NEAR-STATIONARY REGION.

When a near-stationary region has been reached, the immediate neighbourhood is explored by performing a number of suitably chosen experiments which will allow a response function of second degree to be fitted, the equation being

$$Y = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k}\sum_{j=1}^{k} b_{ij} x_i x_j. \qquad \text{------ (4)}$$

This equation represents a system of hyperquadrics but as it stands it conveys very little of the nature of the surface. A better appreciation of what the surface looks like can be obtained by reducing the equation to its canonical form. This is done by shifting the origin to the stationary point, S, and rotating the coordinate axes so that they correspond to the principle axes of the quadrics. The above equation then reduces to

$$Y - Y_S = \sum_{i=1}^{k} B_{ii} X_i^2. \qquad \text{------ (5)}$$

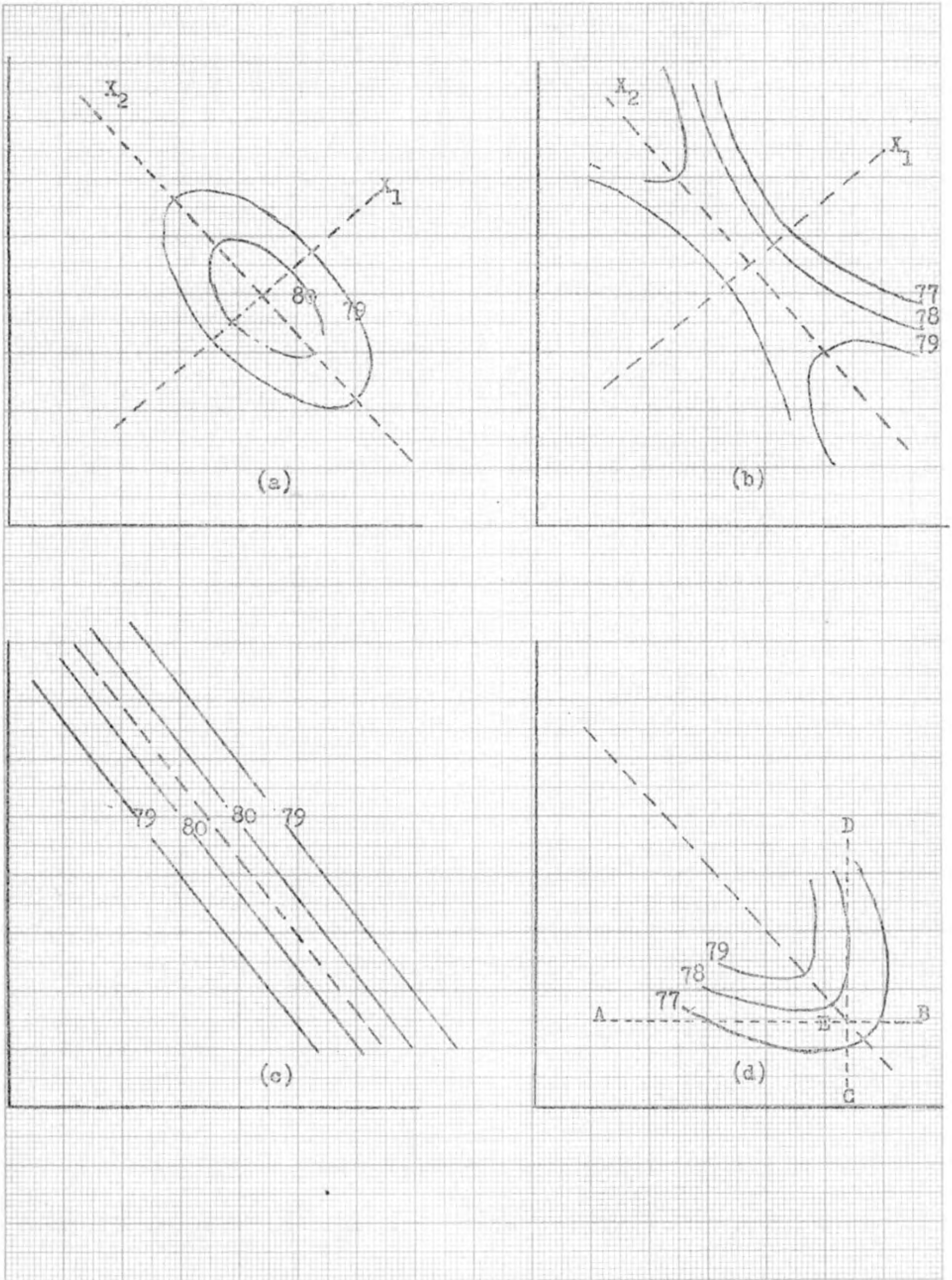
The derivation of the canonical equation is straightforward. The stationary point, S, is found by solving

$$PS = Q \qquad \text{-----} (6)$$

where P is the (kxk) matrix whose elements p_{ij} are equal to $\frac{1}{2}b_{ij}(i \neq j)$ or $b_{ij}(i = j)$. S is the (kx1) matrix of the unknown coordinates of the centre and Q is the (kx1) vector whose elements $q_i = -\frac{1}{2}b_i$. The coefficients B_{ii} are the eigenvalues of matrix P and the direction cosines of the new axes X_i are the corresponding unit latent vectors.

When the fitted surface contains only two independent variables a visual insight into the form of the surface can be obtained by plotting contours of equal response. Such a representation allows the experimenter to study the effect of changing the levels of the factors without having the distraction of representing the response in another direction. Examples of surfaces generated by a second degree equation in two variables are given in Figure 1.1, and these are the only possible types of surface for a second degree equation in two variables. If B_{11} and B_{22} are of the same sign the contours are elliptical with a point maximum (or minimum) at the centre S. If the coefficients are of different signs the contours are hyperbolas and the resulting configuration is known as a saddle point, col, or minimax. The degree of attenuation along the axis depends on the relative magnitudes of the coefficients and in Figure 1.1(b) B_{22} is smaller than B_{11}. When B_{22} is zero we have a stationary ridge system in which the contours are all straight lines, except when the centre of the system is at infinity in which case the contours of the ridge are parabolas. In Figure 1.1, (c) and (d) represent limiting forms of (a) and (b), for example, (c) represents the momentary situation between (a) and (b) when the sign of B_{22}

SURFACES GENERATED BY A SECOND DEGREE EQUATION IN TWO DIMENSIONS.



changes from negative to positive. These limiting forms rarely occur exactly in practice but often we have an attenuated form of one of the basic surfaces (a) and (b) which approaches one of the limiting cases.

When one coefficient is small compared with the other in the canonical equation, but the centre is remote from the design, a ridge of some kind is indicated and the effect of changes in response along this ridge can best be appreciated by making a further transformation of the response equation. Suppose B_{22} is small. Since all large first order effects have been eliminated the axis, X_2, corresponding to the coefficient B_{22}, will normally be found to pass close to the centre of the design. A new origin is then taken on X_2 near to the centre of the design, the co-ordinates of which are X_1, and $X_2' = X_2 - a$. Substituting back into the canonical equation we have

$$Y - Y_s' = B_{11}X_1^2 + B_{22}X_2'^2 + B_2' X_2' \qquad \text{----- (7)}$$

where Y_s' is the predicted value of the response at the new origin,

$$Y_s' = Y_s + a^2 B_{22} \qquad \text{-----(8)}$$

and

$$B_2' = 2a\, B_{22} \qquad \text{----- (9)}$$

is the slope of the ridge X_2 at S'. B_2' is therefore the predicted increase in response for unit changes along the ridge.

For the general case in which there are k factors the interpretation of the response surface follows the principles outlined above for the case of two factors. However it is only permissible to make inferences about the shape of the surface over a limited region which is dependent on the range of the experimental design. Box and Wilson suggested that interpolation should be restricted to a hypersphere whose centre is at the middle of the experimental design whose radius, R, is equal to $\sum_{i=1}^{k} x_i^2$. When there are several independent variables the point at which the response is maximal within the hypersphere is not readily estimated, even when the response equation has been reduced to its canonical form.

This particular problem was dealt with by Hoerl (2) who considered the quadratic response surface whose equation is

$$y = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k}\sum_{j=1}^{k} b_{ij} x_i x_j. \qquad \text{----- (10)}$$

and the hypersphere of radius R $(=\Sigma x_i^2)$, centred at the middle of the design. He developed a method of calculating the paths of maximum and minimum response from the centre of the design to a point on this hypersphere.*
Substituting

$$x_k = \left(R - \sum_{i=1}^{k-1} x_i^2\right)^{\frac{1}{2}} \qquad \text{----- (11)}$$

in equation (10) he then obtained (k-1) equations in x_i(i=1 ....$\overline{k-1}$) and R by differentiating the response equation with respect to x_i, and equating to zero. Substituting back for R and putting

$$\mu = \frac{b_k + \sum_{i=1}^{k-1} b_{ik} x_i}{x_n} \qquad \text{----- (12)}$$

the following equations were derived.

$$\begin{array}{llll}
[2(b_{11} - b_{kk}) - \mu]\, x_1 + & b_{12} x_2 & + \cdots\cdots + b_{1k} x_k & = -b_1 \\
b_{12} x_1 & + [2(b_{22} - b_{kk}) - \mu]\, x_2 + & \cdots\cdots + b_{2k} x_k & = -b_2 \\
\vdots & & \vdots & \vdots \\
b_{1k} x_1 & + \; b_{2k} x_2 & + \cdots\cdots + \mu x_k & = -b_k
\end{array} \qquad \text{----- (13)}$$

* The examiners have pointed out that this problem can be dealt with more neatly with a Lagrange Multiplier.

In an obvious matrix notation this may be written as

$$[H - \mu I]X = G \qquad \text{----- (14)}$$

Hoerl then states, without proof, that the maximum value of y on any radius R is defined by the x_i corresponding to a value of μ greater than the maximum latent root of H. Similarly the absolute minimum is given by x_i corresponding to a value of μ less than the minimum latent root of H. Thus the maximum, or minimum, value of the response on the surface of a hypersphere can be found from

$$X = [H - \mu I]^{-1} G \qquad \text{----- (15)}$$

This is a useful technique when there are more than two factors to consider as it is possible to represent the values of x_i on the optimum ridge from the centre of the design on a two-dimensional graph.

Although Hoerl does not point it out there is a connection with the canonical form discussed previously as the matrix H considered above is related to the matrix P in equation (6), for

$$H = 2(P - b_{kk} I) \qquad \text{----- (16)}$$

As the canonical form of the response equation will generally be calculated in addition to applying Hoerl's technique it is unnecessary to calculate the maximum and minimum latent roots of H as they are equal to twice the maximum, or minimum, coefficient in the canonical equation, minus 2 b_{kk}.

ADDITIONAL EXPERIMENTS

Having determined the approximate form of the response surface in the immediate neighbourhood of a stationary point the next step is to carry out some confirmatory experiments in order to get more precise information on the characteristics of the surface. If a point maximum is indicated these experiments will generally be performed at points along the canonical axes. If a ridge system is indicated the new experiments will be carried out in the direction of the ridges. New estimates of the coefficients in the response equation will then have to be determined including the information from these additional trials. The matrix C^{-1} is known for the basic design and in order to avoid having to recalculate the precision matrix for all the experiments from the beginning, formulae due to Plackett (3), (1), can be used which require only the inversion of a matrix of order N_2, the number of new observations. Besides C^{-1}, the vector B of the estimates of the elements of β and the residual sum of squares, S, are known for the basic design. If the vector of the additional responses is Z, and W is the corresponding matrix of independent variables, then Plackett showed that the new matrices C_1^{-1}, B_1 and S_1 for all the experimental points are given by

$$C_1^{-1} = C^{-1} - J'GJ$$

$$B_1' = B' + V'GV$$

$$S_1 = S + V'GV$$

Where

$$J = WC^{-1};\ G = (I + R)^{-1};\ R = WC^{-1}W';\ V = Z - WB$$

The major saving in computation is in the calculation of the matrix C_1^{-1}, but if the number of additional points is comparable to the initial number of points then the use of Placketts technique will not result in such a great saving as some time will be spent on the additional auxiliary matrix computations. Generally the number of confirmatory trials will be relatively small and the use of Placketts formulae should result in a real saving in computational labour.

PRECISION OF THE ESTIMATED OPTIMUM

Little was said in the original Box and Wilson paper about the precision of the estimated optimum working conditions. In a later paper Box and Hunter (4) determined a confidence region for a stationary point by considering the general problem of finding the limits of the solutions of a set of simultaneous linear equations. The particular equations of interest in quadratic response surfaces are those given by equation (6) and Box and Hunter showed that the confidence region of the solution depends upon the magnitude of the errors in the coefficients, b_{ij}, which in turn depend upon the experimental design and the estimate of the experimental error. It is also shown that the confidence region reflects the general characteristics of the response surface itself.

This particular work has its greatest value when the surface has an absolute maximum, which will be located at the stationary point. However, when the summit is fairly flat and the curvatures small, the exact maximum will be difficult to estimate accurately and the confidence region will be wide. In such a situation as this it is probably not worth devoting a lot of effort to locating the exact maximum as any inaccuracies in its estimation will not result in any serious loss in response due to the summit being so flat. Hartley in the discussion in (1), suggested that it may be preferable to estimate the amount of the maximum response, which is relatively simple, rather than the position. If the estimated maximum is only a little better than the response already obtained then there is no need to proceed further. * See facing page.

1.3 SOME GENERAL CONSIDERATIONS

The general methods of analysis of response surfaces, together with illustrative practical applications (mainly chemical) have been described by Box and Wilson (1), Box (5), Box and Youle (6), Read (7), and Davies (8). These papers are particularly useful to the applied statistician as they examine the various characteristics of response surfaces, how they arise and how they should be interpreted; and it is shown in (6) how it may be possible to determine the basic mechanism of a system from the study of the response surface. Other topics discussed include the joint optimising of several responses, the use of electronic computers in analysing responses surfaces and methods of presenting the results back to the experimenter.

FACTOR DEPENDENCE

The reason for quadratic surfaces having the forms shown in Fig.1.1, is clearly due to the response function for one factor being dependent on the remaining factors. As the factors jointly influence the response it is of little value to vary one factor at a time, especially when the response surface contains a ridge. Consider a surface like that in Fig.1.1.(d) and suppose the experimenter starts at A and varies x_1, keeping x_2 constant. He will reach an optimum at E and if he then keeps x_1 constant at this point and varies x_2 he will find that by moving away from E the response decreases and he will falsely conclude that E is the point of maximum response. For response surfaces of type (a) or (b) in Fig. 1.1 the true optimum may eventually be reached by the one factor at a time approach, but the path followed would be very erratic.

It is pointed out by Box (5) that factors like pressure, time, temperature etc., can only be regarded as "natural" variables because they can be conveniently measured separately but in practice the behaviour of systems can often be described more economically in terms of "fundamental" or "compound" variables which will be functions of one or more of the natural variables. For this reason many combinations of natural variables may correspond to the best level of the fundamental variables and in addition it can happen that a system dependent upon K natural variables can be expressed in L ($<$ K) compound variables. In such situations some coefficients in the canonical form are zero, or very small, and some sort of ridge system is indicated.

An important feature of the discovery of factor dependence is that it may lead to a better understanding of the basic mechanism of the system. Box states in (1), (5), (6) that from his experience in these problems point maxima are something of a rarity and that ridges, stationary or rising, are fairly common. He suggested that the nature of a ridge system could indicate physical laws which underlay the process studied. He discusses this in some detail in (6) and he points out that the compound variables mentioned above can have a greater significance than a purely representational one. This is illustrated in the text of the last reference by both hypothetical and real problems.

SECONDARY RESPONSES

When the surface contains a ridge such as that in Fig.1.1(c), a whole range of optimum conditions exist corresponding to points along the crest of the ridge. The factors are therefore compensating as the effects of changes in one factor can be balanced by changes in one or more of the other factors. This situation can be exploited to great advantage when there is a secondary response to consider for it is then possible to choose that point on the ridge for the main response for which the secondary response is maximised. Thus both the major response and the auxiliary response might be brought simultaneously to their best levels. Alternatively, there may not be a secondary response but other practical considerations can be taken into account. For example, it may happen that some of the alternative conditions on the ridge are less costly or more convenient than others.

PRESENTATION OF RESULTS

When the response surface has been analysed statistically it is necessary to present the results back to the experimenter in a form which is unambiguous and readily understandable. When there are only two or three factors the simplest representation is the geometrical one, particularly when there are several responses to consider. Contour representations in up to three variables are readily constructed and in three-dimensions it is often useful to make a model in which contours can be denoted by coloured wires supported by wire grids. When there are two responses to consider they can be shown on the same model, or two models can be built and viewed side by side. Another means of diagramatic representation is the use of tri-coordinate diagrams.

With more than three factors it is not possible to represent features of the surface geometrically but the coordinates of the ridges of maximum and minimum responses can be graphically represented after applying Hoerl's technique (2). Box in (5) also shows that it is sometimes possible to illustrate ridge systems with 5 variables by listing alternative processes, over the neighbourhood of the optimum, in tabular form.

USE OF ELECTRONIC COMPUTERS IN THE ANALYSIS OF RESPONSE SURFACES

In a second order approximating polynomial for k variables there are $\binom{k+2}{2}$ constants to estimate from the $\binom{k+2}{2}$ linear equations obtained by the application of the method of Least Squares. With orthogonal designs these equations are simple and easily solved on a desk calculator, but often the designs which may be used are non-orthogonal. The siting of the experimental points tends to be dictated by the features of the response surface in the later stages of experimentation and an electronic computer is then required to solve the least squares equations.

It is pointed out in (9) that the analysis of response surfaces is readily carried out on an electronic computer, particularly as most of the calculations can be expressed very simply in matrix notation. The only input necessary is the design matrix, D, the response vector, Y, and parameters specifying the dimensions of the experiment. The X matrix can be generated from the design matrix and subsequent calculations are straightforward. The output will usually consist of B, C^{-1}, the vector of predicted values, XB, the residuals (which will give some check on the validity of the model and on gross blunders in the data), the total sum of squares, the residual sum of squares and the coefficients and axes of the canonical form. Secondary responses can be automatically read in by the programme and the appropriate output obtained; it is worth noting that in such situations the transforming matrix T and the precision matrix C^{-1} are already available within the computer and need not be recalculated. The use of Hoerl's technique and Plackett's formulae can also be readily incorporated into a general programme.

1.4. EXPERIMENTAL DESIGNS FOR RESPONSE SURFACES

DEVELOPMENT OF EXPERIMENTAL DESIGNS

The main requirements of an efficient design for investigating response surfaces are that it should allow the graduating polynomial to represent the true function accurately within a given region, without containing an excessive number of experimental points. It should allow a check to be made on the model, and if the model is found to be inadequate the design should be readily augmented with further points to allow higher order polynomials to be fitted. Finally, the design should lend itself to "blocking" arrangements; that is to say, it should allow the elimination of systematic effects, such as polynomial time trends or block effects, without affecting the precision of the regression coefficients.

The designs employed in the original Box and Wilson paper (1) had mainly been of the conventional type. For fitting equations of the first degree the two level factorial and fractional factorial designs had frequently been used, while for second degree equations the three level factorials had been used and a new "composite" design had beeen developed empirically, consisting of two-level factorial designs with extra points added. Most of these designs satisfy many of the requirements listed above, they are convenient to use in practice, and the analysis of the results is extremely simple.

A more fundamental investigation of first order designs was given by Box (10). He showed that maximum precision of the estimated linear regression coefficients was obtained from orthogonal designs, which include the two-level factorial designs and the multi-factorial designs of Plackett and Burman (11). It was also shown that the orientation of the designs does not affect this property of minimum variance. This means that the design can be rotated so that systematic effects are eliminated without loss of efficiency, and possible bias may be reduced.

Orthogonal second order designs of a sort can be obtained by redefining the independent variables in terms of orthogonal polynomials. Box and Hunter (15) show that this condition of orthogonality refers only to a particular orientation of the design and is lost on rotation of the design. This led them to consider the joint accuracy of the coefficients and they developed some designs of first and second order for which the estimated response had constant variance at all points equidistant from the centre of the design. These are called rotatable designs. When such arrangements are rotated about the centre the variances and covariances of the regression coefficients remain constant. Bose and Draper (16) extended this work and derived general methods for obtaining second order designs for three factors and Gardiner, Grandage and Hader (17) derived new third order designs.

The discrepancies between the graduating polynomial and the true function can occur because of sampling error and the inadequacy of the graduating polynomial to represent the true function. Box and Draper (18) refer to these as the "variance error" and "bias error" respectively. They examined the relative importance of these two sources of error in first order designs and they show that practical designs are similar to those which would be appropriate if the experimental error was ignored altogether.

* See facing page.

THEORY OF ALIASES

In fitting a response equation of order d the estimates of the regression coefficients differ from the true values on account of experimental errors in the response and biases arising when it is impossible to represent the function by an equation of the type fitted. This was recognised by Box and Wilson (1) who suggested that an experimental design should be judged on the apparent precision of the estimated constants and the magnitude of the possible biases in the estimates. The nature of the biases can be determined for any experimental design, for suppose the response function is represented by an equation involving L constants when M (> L) are required for an exact representation. Then it is wrongly assumed that

$$\eta = X_1 B_1 \qquad \text{----- (17)}$$

when, in fact

$$\eta = X_1 B_1 + X_2 B_2, \qquad \text{----- (18)}$$

where X_2 is the matrix of independent variables which have been wrongly omitted and B_2 is the corresponding vector of constants in the response function. The least squares estimates, assuming equation (17) is true, are

$$B_1 = (X_1' X_1)^{-1} X_1' Y \qquad \text{----- (19)}$$

and these in general will be biased since

$$E(B_1) = (X_1' X_1)^{-1} X_1' \eta$$

$$= (X_1' X_1)^{-1} X_1' X_1 \beta_1 + (X_1' X_1)^{-1} X_1' X_2 \beta_2 \qquad \text{----- (20)}$$

so that

$$B_1 \rightarrow \beta_1 + A\beta_2, \quad \text{where } A = (X_1' X)^{-1} X_1' X_2$$

The arrow in this expression indicates that the quantity on the left is an unbiased estimate of the quantity on the right. A is called the alias matrix and the elements, a_{ij} determine the extent of the biases and will depend upon the experimental design. If a particular a_{ij} is zero then b_i

will not be biased by β_j. In general, if the number of constants necessary to represent the function is greater than the number of experiments then the extra constants will either bias the estimates or else appear in the residual mean square.

* See facing page.

STANDARDIZED VARIABLES

The precision matrix, C^{-1}, and the alias matrix, A, depend on the arrangement of the experimental points and therefore supply an objective basis for comparing the relative efficiencies of designs. However both matrices depend upon scale factors and in order to overcome this Box and Wilson proposed that the variables should be standardized by making the marginal second moments of the design configuration the same for all factors. For the i^{th} variable this is defined as

$$S_i^2 = \frac{1}{N} \sum_{j=1}^{N} (X_{ij} - \bar{X}_i)^2, \qquad \text{----- (21)}$$

so that we have the standardized $x_{ij} = (X_{ij} - \bar{X}_i)/S_i$, and in addition

$$\sum_{j=1}^{N} x_{ij} = 0 \qquad \text{----- (22)}$$

$$\sum_{j=1}^{N} x_{ij}^2 = N \qquad \text{----- (23)}$$

DESIGNS OF TYPES A AND TYPE B.

Box and Wilson defined designs of Type A and order d as those which give unbiased estimates of all derivatives of order 1 to d, providing the assumption is true that all terms of higher degree can be ignored. In such designs the number of experiments can be as small as the number of constants estimated. Designs of Type B provide unbiased estimates of all derivatives of order 1 to d even though terms of order d + 1 exist so that the number of experiments must be greater than the number of constants estimated in order that the aliases can be accommodated by the residual degrees of freedom. It should be noted that it is unnecessary to know the value of β_o in order to apply the method of steepest ascent or to study the nature of the response surface so that any bias in b_o is immaterial.

First order designs can be obtained by varying the factors at two levels and factorial designs, fractional factorial designs and the multifactorial designs of Plackett and Burman (11) were considered particularly appropriate. Multifactorial designs are given for k=3,7, ...4m -1, ...99 factors in N = 4,8, ...4m ....100 experiments and provide designs of Type A. When N is a power of two these designs are identical with fractional factorials while for intermediate values of k the next higher designs is used, with the appropriate number of columns omitted from the design matrix. First order designs of Type B can be obtained by duplicating the appropriate Plackett and Burman design of Type A and order 1 with reversed signs, thereby allowing a check to be made on the assumptions, and in addition any factor held constant in the first design can be simultaneously introduced by changing its level in the duplicated design. All these designs are orthogonal with $X'X = NI$ and it can be shown that while first order terms may be biased by the cross-product coefficients β_{ij} they are always free of quadratic effects β_{ii}.

Fractional factorial designs of Type A and order 1 are found by associating independent variables with interaction terms in the complete design and this is illustrated by a design for 7 factors in 8 experiments in which the four "additional" variables are equated to the interaction terms in the basic 2^3 design. When the number of factors is between 3 and 7 the remaining factors should be treated as dummy variables; the aliases are obtained from the alias matrix omitting all terms containing dummy suffixes. Care must be taken in these circumstances since all such designs are not equally satisfactory.

Several factorial designs of Type B are derived in (1) and it is shown, for example, that for 8 factors in 16 experiments a suitable design can be obtained by duplicating with reversed sign the design for 7 factors in 8 experiments and associating x_8 with x_0 in the original design. This particular design allows unbiased estimates of both first order effects and mixed derivatives and such a design is said to be of Type B' and order 1.

The designs of Type B and order 1 discussed above allow unbiased estimates of first order effects to be obtained but to do this it is necessary to have at least twice as many experiments as the number of constants to be estimated and this would appear to be wasteful if linear effects were dominant. Daniel (12) suggested that in situations where the experiments can be performed sequentially a more efficient procedure was possible by varying the factors one at a time. The estimated linear effects will be of minimal precision but if the effects are large then a move can be made towards a higher response relatively quickly. If the effects are not large additional runs can be added to complete the $\frac{1}{2}$-replicate design, thereby maximising the accuracy of estimated linear coefficients and at the same time estimating second order terms. For example, with 3 factors in 4 experiments he proposes to carry out the trials (1), a,b, and c, which will allow unbiased estimates of the main effects to be calculated even if second order terms are important. The precision of the estimates is only half that obtained from the trials (1), ab, ac, bc, but in this latter case linear effects are aliased with second order terms. If the linear effects are found to be small the $\frac{1}{2}$-replicate defined by I = ABC can be completed by adding the run abc so that maximum precision of the linear estimates is obtained and the run (1) can be used in the contrast (a+b+c-abc-2(1))/4 to get an estimate of -AB-AC-BC+ABC.

These one factor at a time sets can only be augmented to $\frac{1}{2}$-replicates and it is unlikely therefore that this procedure will be used for more than 5 factors. In addition this technique violates the requirements of randomization so that the results are of limited validity.

COMPOSITE DESIGNS

Two-level factorial designs are very useful for estimating the constants in an equation of first degree but for second order equations the 3-level factorials are much less useful. This is mainly because of the large number of experiments required relative to the number of constants to be calculated, which would be very inefficient if the experimental error was small. In addition there is also the disadvantage that the estimated constants are of very different accuracies, and in particular the variance of the quadratic coefficients is twice the variance of the interaction derivatives.

The linear and all cross-product coefficients can, however, be determined from 2-level factorial designs of Type B' and Box and Wilson suggested that the additional quadratic terms could be estimated by adding to the basic factorial design a central point and axial points a distance α from the centre of the design, the distribution of the axial points being symmetrical with respect to each factor. Such a design is called a Central Composite design. The value of α can be chosen to make the design orthogonal or alternatively so that all second order coefficients are estimated with equal precision. The precision of the estimated coefficients and the extent of the possible biases indicate that these designs compare favourably with 3-level factorial experiments when the reduced number of trials is taken into account.

These designs have the advantage that they can be carried out in stages. If in the initial factorial or incomplete factorial design, large first order effects are found the method of steepest ascent can be employed, but if the cross-product terms are of comparable magnitude then it will be necessary to determine the quadratic effects and the required points are readily included into the composite design. In situations where the initial design has indicated the existence of a stationary point in a particular direction away from the centre of the design the additional points should be added in that direction thus forming a non-central composite design.

The problem of the possible existence of "blocking" effects when adding the cross-polytope, or radial points, to the original design was considered by De Baun (13). The b_i, b_{ij}, will in general be unaffected by blocking effects but the b_{ii} will be biased and the extent of this bias will depend on the choice of the distance, α, of the radial points from the centre of the design and the shift in b_o. However if the block effect is only to shift b_o, the nature of the response surface will remain unaffected. For the initial Type B' design consisting of N_1 points, n_1, of which are central, the expected value of b_o is given by

$$E(b_o) = \beta_o + (N_1-n_1)/N_1 . \Sigma \beta^2_{ii} \qquad \text{-----} (24)$$

For the cross-polytope of N_2 points the expected value of b_o is

$$E(b_o) = \beta_o + 2\,\alpha^2/N_2\, \Sigma\, \beta_{ii}^2 \qquad \text{-----} (25)$$

For the block contrast to be orthogonal these expectations must be equal so we have to choose α such that the equation

$$2\,\alpha^2 = N_2/N_1(N_1 - n_1) \qquad \text{-----} (26)$$

is satisfied.

The basic design forming the nucleus of the composite design usually consists of a factorial experiment. As the number of factors increases the number of experimental points increases rapidly and for this reason the complete factorial is often replaced by a fractional replicate. The choice of these basic factorial experiments is discussed by Hartley (14) and in particular those designs which do not allow the estimation of all the cross-product coefficients. In such situations he shows that if no main effect is used as a defining contrast in the fractional replicate it is possible to estimate from the composite design all linear coefficients, b_i, all quadratic coefficients, b_{ii}, the constant term, b_o, and one, and only one, of the product terms, b_{ij}, from each of the alias sets. This implies that the largest number of coefficients can be estimated if each alias set contains at least one two-factor interaction, $x_i x_j$. He examines some of the usually recommended fractional replicates, for example Davies (8), which keep main effects clear of two-factor interactions but which do not permit the estimation of all individual interaction terms, and shows that by changing the defining contrasts and aliasing linear and/or product terms in the basic design it is possible to get information on more coefficients from a composite design. He discusses the efficiencies of these alternative designs and shows that there is a loss in the precision of a linear coefficient through introducing a product term into the alias set.

It is doubtful if the above procedures are of real practical benefit in the study of quadratic response surfaces as it is generally necessary to estimate all the second degree coefficients to represent the surface adequately. If the experimental design does not allow the estimation of all the necessary coefficients then some of the coefficients which are estimated will be biased and a true representation of the surface will not be possible. While Hartley's proposals may be of use in other situations they hardly seem applicable to the present problem.

FIRST ORDER DESIGNS OF MAXIMUM PRECISION

The problem of choosing first order designs for which the variance of each of the linear coefficients is minimised was considered by Box in (10). The matrix of variances and covariances for the estimates of the regression coefficients is $(X'X)^{-1}\sigma^2$ and the problem is therefore to choose the design matrix so that the diagonal elements of $(X'X)^{-1}\sigma^2$ are minimised. When the standardised variables are functionally independent Box showed that this condition is satisfied only when all the off-diagonal terms in the precision matrix are zero so that the linear regression coefficients are uncorrelated and have variance σ^2/N. Maximum efficiency is therefore obtained from an orthogonal design. This is a generalisation of the result obtained by Plackett and Burman (11) who had restricted the standardised values of the independent variables to $\pm$ 1.

Designs of optimum precision for up to $k = N-1$ factors in N experiments can be obtained from any orthogonal matrix with all elements in the first column equal. When $k = N-1$, the geometrical implication of this result was shown to be that the N experimental points are at the vertices of a regular $N-1$ dimensional simplex and if $k < N-1$ the experimental points are the projections onto a space of k dimensions of the vertices of the $N-1$ dimensional regular simplex. No restriction is imposed on the orientation of the design as this does not affect the variance-covariance matrix but corresponds to a different choice of the orthogonal design matrix. This class of designs includes the factorial experiments which, of course, are of particular value as in addition they readily allow a check on the adequacy of the first degree equation, they form a good basis from which designs of higher order may be augmented and they lend themselves to blocking.

While the variances and covariances remain constant under orthogonal rotation of the design the biases which may occur, due to the planar representation being inadequate, do not. As a measure of the effect of bias Box considered the sum of squares of bias coefficients for the $k + 1$ coefficients in the equation of the plane. These are given by the diagonal terms of AA', where A is the alias matrix, but it was found that this matrix is invariant under orthogonal rotation of these optimum first order designs. He then concluded that no particular design can be chosen in preference to another if no prior knowledge was available concerning the relative importance of the second order terms. However if the direction of the principal axes of the canonical form were known then by choosing the design so that the axes of the design were parallel to the principle axes, the linear coefficients would be unbiased, since in the new variables the second degree equation contains no product terms and the quadratic terms could be aliased with b_o.

These orthogonal designs for estimating linear coefficients are also very useful for eliminating systematic variations such as block effects and time trends. It has been shown that first order designs of maximum efficiency can be obtained by taking k mutually orthogonal column vectors, each containing N elements, and each orthogonal to a column vector of unit elements. It is well known that another N-k-1 vectors can be found which are orthogonal to these k-1 vectors and which are mutually orthogonal; consequantly some of these N-k-1 vectors can be used to represent the systematic variations.

ROTATABLE DESIGNS

The criterion of orthogonality has been found very useful in determining optimum designs of first order. For designs of order higher than the first the quantities $x_o, x_1 \ldots x_k, x_1 x_2, \ldots x_1^2, \ldots$ are not functionally independent and a diagonal moment matrix $N^{-1}(X'X)$ is not possible. Orthogonal designs of a sort can be obtained by redefining the independent variables in terms of orthogonal polynomials but it is shown by Box and Hunter (15), that there is an infinite variety of such designs with widely different properties; examples of such designs are 3-level factorials and orthogonal composite designs. They considered the moment matrix and showed that several of its elements would have to be made zero in order to have a diagonal matrix but for some moments corresponding to diagonal terms, the choice is arbitrary. The choice of these terms is discussed at some length and it is shown that they have conflicting effects for in choosing them such that the variance of quadratic effects is reduced they cause an increase in the possible bias in the linear effects.

It is then demonstrated that while the curvature in one direction may be measured with a certain precision the accuracy with which it is determined in another direction, which may be of equal importance, may in fact be much reduced. In the 3-level factorials, for example, the covariances between all effects are zero in the normal orientation of the design but if the design is rotated the variances and covariances of the second order effects undergo marked changes and only the linear effects have constant variances in all orientations. Thus the variances of individual coefficients in a particular orientation may give a misleading impression of the efficiency of a design and the condition of orthogonality refers to a particular orientation and is lost on rotation.

Considering the accuracy of individual coefficients does not lead to any unique solution for designs of order higher than the first, so Box and Hunter considered the joint accuracy of the coefficients. The variance of the estimated response at some point x is a function of the independent variables, σ^2 and N so that $V(x) = NV(y_x)/\sigma^2$ is a unique standardization of the accuracy with which a response at x is estimated, and this is called the variance function of the design. It was suggested that if nothing was known of the orientation of the surface, as is usually the case, it would seem appropriate to choose designs for which the contours of the variance functions are hyperspheres centred at the origin of the design so that the response is estimated with constant precision at all points equidistant from the centre of the design. An arrangement of points having this special variance function is called a rotatable design. * See facing page.

Conditions for rotatability were derived and it was found that a necessary condition was that the moment matrix should be invariant under orthogonal transformation of the design matrix. The moments which are elements of this matrix are the same as those of a spherical distribution and are known apart from constants $\lambda_0, \lambda_2, \ldots \lambda_{2d}$. Writing $(1^{\alpha_1}, 2^{\alpha_2}, \ldots k^{\alpha_k})$ for the moment $N^{-1} \sum_{u=1}^{N} x_{1u}^{\alpha_1}, x_{2u}^{\alpha_2}, \ldots x_{ku}^{\alpha_k}$ it is shown that when any of the α_i are odd

$$(1^{\alpha_1}, 2^{\alpha_2}, \ldots k^{\alpha_k}) = 0, \qquad \text{-----} (27)$$

and when all α_i are even

$$(1^{\alpha_1}, 2^{\alpha_2}, \ldots k^{\alpha_k}) = \frac{\lambda_\alpha \prod_{i=1}^{k} \alpha_i!}{2^{\alpha/2} \prod_{i=1}^{k} (\frac{1}{2}\alpha_i)!} \qquad \text{-----} (28)$$

The variance function therefore depends only on λ and $\rho = (x'x)^{\frac{1}{2}}$ so the required form of the design matrix is completely defined by choosing the λ's to give a suitable variance matrix and alias matrix.

From equation (28), $\lambda_0 = [0]$ and $\lambda_2 = [11]$ and since by convention $[0] = [11] = 1$, λ_0 and λ_2 are always equal to unity. There is therefore no choice for the λ's in a rotatable design of first order for which the moment matrix is the identity matrix, so that the condition for rotability is precisely the same as that it should have minimum variance. For second order rotatable designs the linear and product coefficients are necessarily uncorrelated but the correlation between the quadratic terms depend upon the choice of λ_4. If λ_4 is made equal to unity these correlations are zero and the design is therefore orthogonal as well as rotatable. In such a situation it is worth noting that the variances of the quadratic terms are only half the variances of the product terms so that four times as much weight is placed on the quadratic terms relative to the product terms in these designs compared with 3-level factorials. With these orthogonal designs the precision is high particularly near the centre of the design, but the bias coefficients are also high. With slightly lower values of λ_4 a more uniform distribution of precision is obtained over the immediate vicinity of the design, and in addition the third order bias coefficients are reduced. No clear cut statement is made on the best choice of λ_4 and it was suggested that a compromise be sought between experimental error and bias which will be satisfactory for practical situations.

Since the moments of a rotatable design of order d must be the same as the moments up to 2d of a spherical density function it seemed reasonable to assume that rotatable designs might be constructed by equally spacing the experimental points on one or more hyperspheres. A set of points all equidistant from the centre of the design are called equiradial sets and if the moments up to 2d are such that they are unaffected by rotation they are called equiradial rotatable sets of order d. These sets are discussed at length by Box and Hunter who consider individual cases in two dimensions and three dimensions before generalising the problem. It is shown that it is necessary to have more than one radial set otherwise the quadratic coefficients and the constant term cannot be separately estimated. This is often overcome by adding central points, corresponding to a set of radius zero, which in addition provides a useful means of modifying the value of λ_4. In general λ_4, and therefore the variance and bias coefficients is dependent upon the number in the different sets, the radii of the sets and the number of independent variables.

For two dimensional second order rotatable designs each ring of non-zero radius must contain a minimum of five points. Orthogonal designs can be obtained from pentagonal and hexagonal designs with five or six points respectively, and for hexagonal designs there is the added advantage that the constant term and second order coefficients are unbiased by third order effects in any rotation. Rotatable designs with more than ten points can also be obtained by combining two or more circles of equispaced points but in addition it is also possible to combine sets of equiradial points which in themselves are not of order 2; for example, sets of three points forming the vertices of an equilateral triangle with centre coincident with the origin.

In three dimensions, sets of equally spaced points on a sphere are provided by the tetrahedron, octahedron, cube, icosahedron and dodecahedron. Only the latter two, when combined with central points individually, provide second order rotatable designs; they can also be combined. As with two dimensional designs, sets of points not themselves rotatable arrangements of order two can be combined to give them. One particular interesting combination is the cube and the octahedron orientated such that a line joining the origin to the vertex of the cube passes through the centre of the face of the octahedron; these designs are special cases of the central composite design.

In five or more dimensions there are only three regular figures, the regular simplex, the cross polytope and the hypercube. The most useful designs can be obtained by combining the cross-polytope and the hypercube to form particular cases of the composite design. Consequently they lend themselves to sequential testing and are easy to use.

Rotatable designs can conveniently be performed in orthogonal blocking arrangements and Box and Hunter show that the two necessary conditions for this are:-

(1) the sum of products between the independent variable $x_0, x_1 \ldots . x_k$, must be zero for each block,

(2) the fraction of the total sum of squares for each variable in each block must be proportional to the number of observations in each block.

For rotatable designs consisting of equidistant sets of points with added central points the above conditions can be satisfied by dividing the equiradial sets into subsets which are themselves first order rotatable designs and adding the appropriate number of central points to each of these subsets. With composite designs it is not always possible to attain both exact rotatability and orthogonality for quadratic variables and block effects and in practice it is usually more convenient to retain the orthogonality and to relax the conditions of rotatability. In these latter designs both the hypercube and the cross-polytope are first order designs and conveniently define the blocking arrangements and in addition the hypercube may sometimes be further subdivided into first order rotatable designs, for example the fractional factorials.

At this time there were very few rotatable designs known and a fundamental investigation of second order rotatable designs in three dimensions was carried out by Bose and Draper (16). The conditions for rotatability of second order designs for three factors, derived by Box and Hunter, are

$$\sum_{u=1}^{N} x_{1u}^2 = \sum_{u=1}^{N} x_{2u}^2 = \ldots . = \sum_{u=1}^{N} x_{ku}^2 = \lambda_2 N. \qquad \text{------ (29)}$$

$$\sum_{u=1}^{N} x_{1u}^4 = \sum_{u=1}^{N} x_{2u}^4 = \cdots = \sum_{u=1}^{N} x_{ku}^4 = 3\sum_{u=1}^{N} x_{iu}^2 x_{ju}^2 = 3\lambda_4 N \qquad \text{----- (30)}$$

$$\lambda_4 / \lambda_2^2 > k/(k+2) \qquad \text{----- (31)}$$

Condition (31) can always be satisfied by adding central points to the design.

Bose and Draper define certain transformations applied to points in three dimensions as follows. Let $W(x,y,z) = (y,z,x)$. Then $W^2(x,y,z) = (z,x,y)$ and $W^3(x,y,z) = (x,y,z)$. Thus W, W^2 and $W^3 = I$ form a cyclical group of order 3. Let $R_1(x,y,z) = (-x,y,z)$, $R_2(x,y,z) = (x,-y,z)$ and $R_3(x,y,z) = (x,y,-z)$. The four transformations W, R_1, and R_2 and R_3 thus generate a group G of 24 points with coordinates $(\pm x, \pm y, \pm x)$, $(\pm y, \pm z, \pm x)$, $(\pm z, \pm x, \pm y)$. This set $G(x,y,z)$ satisfies the moment conditions (29) but not

$$\sum_u x_{iu}^4 = 3\sum_u x_{iu}^2 x_{ju}^2 \qquad (i \neq j), (i, j = 1,2,3) \qquad \text{----- (32)}$$

The fraction $K(x,y,z)$ of the point (x,y,z) is then defined as

$$K(x,y,z) = 1/3\,(x^4+y^4+z^4-3y^2z^2-3z^2x^2-3x^2y^2) \qquad \text{----- (33)}$$

This function is constant for all 24 points of the set G and if it is equal to zero then $G(x,y,z)$ is a rotatable design. If $K(x,y,z) \neq 0$, then $\sum K(x,y,z)$ over a point set is defined as the excess of the set. Thus the excess over the set G is

$$Ex[G(x,y,z)] = 8(x^4+y^4+z^4-3y^2z^2-3z^2y^2-3x^2y^2), \qquad \text{----- (34)}$$

and this can take positive or negative values depending on x,y and z. . Consider the 12 points $(\pm p, \pm q, 0)$, $(\pm q, 0, \pm p)$, $(0, \pm p, \pm q)$, and denote this set by $\frac{1}{2}G(p,q,0)$ this set has excess

$$Ex[\tfrac{1}{2}G(p,q,0)] = 4(p^4+q^4-3p^2q^2) \qquad \text{----- (35)}$$

The set $\frac{1}{2}G(P,Q,0)$ will itself form a rotatable design if the excess is zero. This implies that $p^2/q^2 = (3 \pm \sqrt{5})/2 = \theta$ and θ^{-1}, say. Thus, the set of points reduces to $(\pm\theta, \pm 1, 0)$, $(\pm 1, 0, \pm\theta)$, $(0, \pm\theta, \pm 1)$. These are the vertices of an icosahedron and by adding central points the icosahedron design of Box and Hunter (15) is obtained.

Bose and Draper then consider the combination of several points sets which in themselves may not form rotatable designs but which when combined have zero excess and thus form rotatable designs. For example, the combination of the two sets $\frac{1}{3}G(a,a,a)$ and $\frac{1}{4}G(c,o,o)$ give the cube plus octahedron design of Box and Hunter (15). In this particular example there are two parameters a and c and as there are two conditions to satisfy, namely $Ex(\text{set})=0$ and $\lambda_2=1$, a unique design is obtained. When there are more than two parameters an infinity of designs are obtained. These are examined and it is shown that previously known designs are special cases of this infinite class.

A second method of generating point sets suitable for building second order rotatable designs is given in terms of polar coordinates. The procedures for obtaining rotatable designs are similar to those given above but in this case the excess of each single point is not constant and it is necessary to consider the total over all points.

The methods were extended to a 16 point design class including the 12 points in the set G for which the product of the coordinates is xyz. The set is written as $G^{(+\frac{1}{2})}(x,y,z)$, and the complementary set for which the product of the coordinates is $-xyz$ is written as $G^{(-\frac{1}{2})}(x,y,z)$. These sets do not in themselves form rotatable designs as they do not satisfy condition (32) nor the further condition that $\sum_u x_{1u}x_{2u}x_{3u} = 0$. Defining a second excess function as

$$Fx[\text{set }(x_{1u}x_{2u}x_{3u}0)] = \sum_u x_{1u}x_{2u}x_{3u}, \qquad \text{-----} \ (36)$$

it was shown that if a set of points S satisfied $Ex(S) = 0$ and $Fx(S) = 0$ then S is a rotatable design of second order. By considering the combined sets $G^{(\pm\frac{1}{2})}(x,y,z)$ and $\frac{1}{2}G^{(\pm\frac{1}{2})}(a,a,a)$ Bose and Draper derive an infinite class of designs using only 16 points.

Third order rotatable designs were investigated by Gardiner, Grandage and Hader (17) who examined combinations of regular and semi-regular figures. They showed that designs for two factors require 7 or more points on each of two concentric circles of different non-zero radii. Each of these sets of points can be rotated independently of the other so there are an infinite number of such arrangements. Points at the centre do not disturb the moment properties of the configuration and can be used to vary λ_4 and λ_6. By a suitable choice of ρ_1, and ρ_2, the radii of the two concentric circles, this design can be carried out in two blocks and the regression coefficients can be estimated independently of block effects. This is particularly useful for generating third order designs from second order designs. It is also shown how a third order rotatable design in two-dimensions may be sequentiallized in 3 stages with a total of 3 or 4 blocks by suitable choice of the number of central points.

Third order rotatable designs in three dimensions are also discussed. Combinations of cubes, octahedra, icosahedra and dodecahedra are examined and it is shown that some of the resulting designs are unsatisfactory as the resulting matrix of normal equations are poorly conditioned. The use of a truncated cube with other figures is studied and sequential methods of developing third order designs are given. This is extended to four dimensions in which four-dimensional analogues of the cube, octahedron and truncated cube are combined. In this latter case 128 points are necessary, not including any central points, and the authors concluded that extension of these principles to higher dimensions was impracticable because of the excessive number of points required.

A BASIS FOR THE SELECTION OF A DESIGN.

In all the previously mentioned papers it had been recognised that discrepancies between the graduating polynomial and the true function occur because of sampling error and the inadequacy of the fitted polynomial to represent the true funtion. No objective basis had been found for comparing the relative importance of the "variance error" and the "bias error" and the designs had been chosen intuitively. This was investigated by Box and Draper (18) who considered the general problem of choosing a design such that

(a) the graduation polynomial most closely represents the true function over a given region, R,

(b) subject to (a), there is a high probability that any inadequacy of the fitted function will be detected.

Denote the variance error by V and the bias error by B. To meet the first requirement the design is chosen to minimise J, the expected mean squared deviation from the true response, averaged over the region R, and normalised with respect to the number of observations and the variance. Thus

$$J = V + B$$

$$= N/\sigma^2 \int_R V[\hat{y}(x)]\, dx \Big/ \int_R \; + N/\sigma^2 \int_R \left[E(\hat{y}(x) - \eta(x)\right]^2 dx \Big/ \int_R \qquad (37)$$

V is minimised by making the design as large as possible, but increasing the size of the design increases the error due to bias. What is required is the best balance between these conflicting issues.

To test for lack of fit the Residual Sum of Squares, S_R, is compared with the experimental error variance. The parameter which determines the power of the test for goodness of fit will be the quantity

$$\sum (E(\hat{y}_u) - \eta_u)^2 = E(S_R) - \nu\sigma^2, \qquad \text{-----} \ (38)$$

where ν is the number of degrees of freedom on which S_R is based. To maximise the power it is necessary to maximise $E(S_R)$, and this is requirement (b).

Box and Draper suggested it was reasonable to regard (a) as being of major importance. The problem is therefore to find a class of designs which satisfy (a) and select from these a sub-class which makes $E(S_R)$ large. They consider in particular the problem of choosing first order designs.

For the case of fitting a straight line for one variable when the true function is more complex it is shown that both requirements (a) and (b) independently require that the third moment of the design points be zero, and in order that the quadratic tendency in the true model be readily detected the fourth moment should not be too small. If there was no bias error the experiment should be as large as possible; if the contributions due to error and bias are equal the size of the design is such that the root mean square distance of the design points from the centre is 0.62θ, with 2θ the width of the region R over which the linear approximation is required. This is very close to the value $0.58\,\theta$ which would be chosen if experimental error were ignored completely.

The general problem of fitting a hyperplane in k variables when the true function is quadratic is considered. The region of interest R, is specified as a hypersphere, defined by $\Sigma x_i^2 \leq 1$. An interesting result of this analysis was that minimising V or B individually required that the design be orthogonal. For minimising total variance, J, the design must be orthogonal with the third moments of the design zero and the variances of the x's equal. The optimum size of the design is such that the root mean square distance of the design points from the origin is greater than $k/k+2$ but less than unity. These are first order orthogonal designs of Type B which were discussed in (1), examples of which are the fractional factorial designs and Plackett and Burman designs.

The authors intend to examine topics such as the effect of cubic bias in first order designs, the extension of the present ideas to second order designs, and the effect of changing the criterion to the minimisation of the maximum mean square error instead of minimisation of average mean square error.

CHAPTER 2.

FACTORS AFFECTING BRITMAG GRADING - No.3 KILN TRIAL

2.1. INTRODUCTION

The initial part of the process at Palliser Works is concerned with precipitating magnesium hydroxide by reacting hydrated dolime with pre-treated sea water. The precipitate is settled, the sludge is filtered on rotary vacuum filters and the residual paste, which contains about 33% in weight of magnesia, is carried on conveyor belts to the feed end of the rotary kilns. The paste is removed from the conveyor belt by a scraper which is made more effective by fine water sprays which keep the face of the belt wet. The paste is fed into the kiln through a gear pump whichhas a variable rate of rotation and can therefore control the rate of feed.

A typical rotary kiln is an inclined revolving steel cylinder approximately 150ft. long by 9 ft. in diamter and lined with refractory brick. The front of the kiln is closed by a hood which carries the burner, through which a mixture of powdered coal and air is blown. This air which acts as a conveyor for the coal is called the Primary air. Ignition of this mixture produces an intensely hot flame. As the kiln rotates the magnesia paste gradually works its way down towards the burning zone and the discharge port. During its passage water is evaporated, alkali salts are volatilized and the residue partially fuses to form the brownish granules known to the refractories industry of this country as Britmag. The size grading of the Britmag granules varies from about $\frac{1}{4}$" in diameter down to very fine dust. The maximum temperature attained in the kiln is in the region of 1650°C. and the exhaust gases leave the kiln at approximately 300°C.

The Britmag is discharged from the kiln into a rotary cooler and is then conveyed to the storage and loading bunkers. The average time taken from the paste entering the kiln to leaving the cooler bin as Britmag is approximately six hours and this time lag had to be taken into consideration when planning the experiment.

2.2. STATEMENT OF THE PROBLEM & METHOD OF ANALYSIS

In this experiment, four factors

(a) the speed of the gear pump (no. of secs. /10 revs): x_1,

(b) the exhaust gas or back end temperature ($^{\circ}$C) : x_2,

(c) the speed of rotation of the kiln (r.p.m.) : x_3,

(d) the volume of primary air (orifice meter inches water gauge differential): x_4,

were varied and their effect on the grading analysis of the Britmag determined. The object of the investigation was to find the levels of these four factors for which the percentage of Britmag less than $\frac{1}{4}$" in diameter and greater than 0.0234" (30 B.S. mesh) in diameter was a maximum. (The percentage of Britmag in this range will henceforth be referred to as the response).

It was proposed to fit a Response Surface of the form

$$\eta = \phi\,(x_1, x_2, x_3, x_4),$$

where η is the true response and ϕ is some function of the variable factors x_1, x_2, x_3, x_4.

Providing this surface is smooth it can be represented to any required degree of approximation by taking a sufficient number of terms in the following series, suitably choosing the constants $\beta_0, \beta_1, \ldots, \beta_{12}$ ... etc.

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{11} x_1^2 + \dots + \beta_{44} x_4^2 + \beta_{12} x_1 x_2 \dots + \beta_{34} x_3 x_4 \dots .$$

Over a sufficiently narrow range, not including a turning point of the surface, a good approximation is given by the plane

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 .$$

If a larger region is to be covered, or if the region is near a turning point it would be necessary to include terms of higher degree which would take into account the curvature of the surface.

As this process had been in operation for some time it was expected that conditions would already be somewhere near the optimum. Consequently it was expected that second order effects would be important and so the experiment was designed to allow first and some second order effects to be calculated. If conditions were already at their best then this investigation would confirm the situation and at the same time give some indication of the effect of changing the levels of the factors. If it turned out that the present conditions were not the best then this analysis would rectify the situation and lead to a more fruitful working point.

Accordingly the four factors were each varied at two levels, forming a single replication of a 2^4 Factorial design, and the sixteen trials carried out in random order to eliminate bias.

To take into account the time lag between the paste entering the kiln and leaving the cooler bin as Britmag, a settling down period of twelve hours duration was allowed between trials. It was suspected that the experimental error, including the effect of un-controllable factors, would be quite large and to offset this seven samples of the Britmag were taken at two-hourly intervals for each trial and the average of these seven samples taken as the observed response for a particular set of conditions. This meant that the testing of each treatment combination took twenty four hours to complete.

2.3. EXPERIMENTAL RESULTS - ANALYSIS & CONCLUSIONS

The levels for x_1, x_2, x_3, x_4 investigated in the experiment are given in Table 2.A.

The factor levels with the corresponding responses obtained are shown in Table 2.B, in the actual order in which the experiment was carried out.

TABLE 2.A.

FACTOR LEVELS FOR THE EXPERIMENT

Factor	Factor Level -1	Factor Level +1
x_1 Gear Pump speed ...... no. of sec. for 10 revs.	86	88
x_2 Back-end Temperature ...... °C.	285	295
x_3 Kiln speed ...... revs. per min.	0.59	0.61
x_4 Primary air ..... "w.g. diff. pressure	0.85	0.95

TABLE 2.B.

EXPERIMENTAL DESIGN AND RESPONSES

TRIAL	x_1	x_2	x_3	x_4	Response (%) y
1	1	-1	-1	1	70.35
2	1	-1	1	1	72.54
3	-1	1	-1	-1	75.19
4	-1	1	1	-1	74.21
5	1	1	1	-1	72.64
6	-1	-1	-1	-1	72.41
7	-1	1	-1	1	67.07
8	-1	-1	1	1	66.33
9	1	-1	-1	-1	62.54
10	-1	-1	1	-1	63.29
11	1	1	1	1	67.08
12	-1	1	1	1	65.29
13	1	1	-1	-1	64.06
14	1	1	-1	1	67.19
15	-1	-1	-1	1	64.21
16	1	-1	1	-1	66.66

Attention may be drawn to the fact that the design of this experiment did not allow quadratic terms of the form $\beta_{ii}x_i^2$ to be determined. This was not considered to be of primary importance at this stage as we were interested in finding out whether second-order effects were comparable to first-order effects and this would be brought to light from a study of the cross-product terms. If this turned out to be the case, supplementary points could be added to the original design which could allow the remaining quadratic terms to be established.

It was intended to calculate an equation of the form,

$$\eta = \beta_0 x_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 \;\ldots\ldots + \beta_{34}x_3x_4 \qquad \text{------ (1)}$$

where x_o is a dummy variable which is always equal to unity.

Owing to the nature of the design the sum of products of any two columns in Table 2.B. is zero and this orthogonality enables the effects to be estimated very simply. It can easily be seen that,

$$b_i = \Sigma yx_i / \Sigma x_i^2,$$

where b_i is the estimate of β_i , and the variance of this estimate is given by

$$\mathrm{Var}\,(b_i) = \sigma^2 / \Sigma x_i^2$$

where σ^2 is the experimental error variance.

As stated earlier seven samples were taken during each trial over a period of twelve hours in order to compensate for sampling and testing errors and for extraneous sources of variation. These latter variations might be due to erratic fuel consumption, for example, or any other uncontrollable factors affecting the process.

There was no evidence of the seven readings within each trial being serially correlated and it was therefore concluded that the sampling and testing error must be much greater than the error due to the extraneous variations in the process. The variation within each trial was consequently used to furnish an estimate of the experimental error. This means that the trials have effectively been replicated seven times. * See footnote.

Although it was not thought of until after the experiment had been carried out it would have been wiser to run a preliminary test in which samples were taken every hour, or even half hour, from a stable process, such measurements being taken continuously for about two days. An examination could then have been made of the correlation between successive readings in a manner described by Jowett, (19). As it turned out nothing was lost by omitting this preliminary study, but it may well have been that the seven results were so highly correlated as to have provided very little increase in efficiency.

There was a particularly odd result in trial 8 and it was omitted from the analysis. The estimate of the residual error variance was calculated from the remainder of the data to be 12.84 and the variance of the mean of seven readings is thus

$$12.84/_7 = 1.834$$

The calculations of the regression coefficients are given in Table 2.C.

* The conclusions are not strictly valid as the extraneous random variations may require a longer period than the 12 hours considered to show themselves. It is later shown that there are no such important extraneous variations and the estimate of the within-trials error can then be used to assess the effects of the factors.

TABLE 2.C.

THE ESTIMATES OF THE CONSTANTS AND REGRESSION ANALYSIS.

Constant Estimated (1)	Σx^2 (2)	Σyx (3)	Estimate =(3)/(2) (4)	Degrees of Freedom (5)	Component SofS =$(3)^2/(2)$ (6)	Mean Square = (6)/(5) (7)
$\hat{\beta}_0$	16	1092.74	68.30	1	74,630.044	
β_1	16	-6.62	-0.414	1	2.739	
β_2	16	12.72	0.795	1	10.112	
β_3	16	6.70	0.419	1	2.806	
β_4	16	-9.26	-0.579	1	5.359	
β_{12}	16	-14.96	-0.935	1	13.988	
β_{13}	16	22.86	1.429	1	32.661	
β_{14}	16	31.78	1.986	1	63.123	
β_{23}	16	4.72	0.295	1	1.392	
β_{24}	16	-29.68	-1.855	1	55.056	
β_{34}	16	1.50	0.094	1	0.141	
		Due to regression		10	187.377	18.7377
		Residual		5	61.982	12.396
		Total		16	74,879.403	

The residual mean square of 12.4 is significantly larger than 1.83, which is the expected value if the postulated second degree surface is correct. The quadratic terms, β_{ii}, are included in the estimated b_o and therefore do not contribute to the residual sum of squares which is made up of second and third order interactions. It appears then that some of third degree terms must be appreciable and the residual was therefore separated into the individual higher order interactions. These are tabulated below.

* See facing page.

TABLE 2.D.

ESTIMATES OF THIRD AND FOURTH ORDER REGRESSION COEFFICIENTS

Constant Estimated (1)	Σx^2 (2)	Σyx (3)	Estimate =(3)/(2) (4)	Degrees of Freedom (5)	Component s.s = $(3)^2/(2)$ (6)
β_{123}	16	-0.40	-0.025	1	0.01
β_{124}	16	-2.56	-0.160	1	0.41
β_{134}	16	-22.74	-1.421	1	32.31
β_{234}	16	-20.48	-1.280	1	26.22
β_{1234}	16	6.96	0.474	1	3.03
			Total	5	61.98

Two of the third degree coefficients are therefore of the same order of magnitude as those of lower degree and it would appear that a second degree model is anadequate. This was rather unexpected as the levels of the factors considered had been selected over rather limited ranges, in fact over narrower ranges than they should have been, possibly, in some cases.

A study of the results obtained in Table 2.B. show that after the sixth trial the responses were all less than 70 per cent whereas the first six trials all yielded more than 70 per cent. This changeover coincided with a report that the kiln was becoming harder to control due to a change in the physical characteristics of the magnesia paste. This was a purely qualitative observation as no measurements of paste properties had been taken.

Whether the properties of the paste had any effect on the grading analysis of the Britmag was at the time open to question. Subsequent laboratory investigations showed that pre-drying the paste and working it mechanically before subjecting it to treatments similar to those experienced in the kiln did in fact strengthen the paste, (see Chapter 3).

As it happened, maintenance of the specified conditions became increasingly difficult until at last experimentation was suspended after the 13th trial due to the inability to control the kiln on a fixed set of conditions for a 24-hour period. This suspension lasted a month or so before the final three trials were able to be completed. It should be pointed out here that each trial does not necessarily follow the previous one on consecutive days as the experiment was sometimes interrupted due to a failure in some part of the process and consequently had to be restarted; but with the exception of the break between trials 13 and 14 no important time lapse occurred between consecutive tests.

It had at one time been considered to treat these last three trials as missing observations and in working on this particular problem a method of estimating missing values in Factorial Designs was derived which is believed to be original, (see Chapter 5).

Recordings of the amount of product leaving the cooler over a period of half a minute had been taken every two hours during each trial. As x_1 measures the number of seconds per 10 revolutions of the gear pump a negative correlation would have been expected to exist between x_1 and the amount of product from the cooler, but a study of the mean production figures for each trial showed that this correlation was quite small. This could be because the two levels of gear pump speed were not far enough apart to show a significant effect or else the gear pump was unsuitable as a measure of the amount of paste fed into the kiln. The former reason is probably the stronger of the two but it must be remembered that the effect of the gear pump is dependent upon the consistency of the paste which is known to vary.

The effect of the primary air is dependent upon the amount of coal used and as there was no means of measuring the coal consumption and therefore making the necessary allowances and it was considered that this could indeed mask any effect of the primary air. Similarly there is a limitation to the levels over which the kiln speed could be varied, due to the mechanical arrangement of the process, and once again it was not sure whether or not extraneous sources of variation would mask its effect. In the case of temperature however the difference between the upper and lower levels, namely 10°C, should be sufficient to show an effect, if the temperature did indeed affect the grading. At this stage it was discovered that the temperature had, in error, been held at the wrong level in trial 8 which naturally invalidates the previous analysis, though it is almost certain that this would not account for the third order coefficients being so large.

The abrupt change in the response between trials 1 - 6 and 7 - 16 was sufficiently marked to justify these two sets being considered as "blocks", especially when there is the practical confirmation regarding the observed change in the initial paste. The effects of some of the factors are thus correlated with these "blocks" and this is illustrated in Table 2.E.

TABLE 2.E.

CORRELATION BETWEEN FACTORS AND "BLOCKS".

Factor	Correlation with "Blocks"	Regression Coefficient
x_1	0	-0.41
x_2	-0.125	0.80
x_3	0	0.42
x_4	-0.25	-0.58
x_{12}	-0.125	-0.94
x_{13}	0.25	1.43
x_{14}	0.50	1.99
x_{23}	0.125	0.30
x_{24}	-0.625	-1.89
x_{34}	0	0.09
x_{123}	0.125	-0.03
x_{124}	-0.125	-0.16
x_{134}	-0.25	-1.42
x_{234}	-0.375	-1.28
x_{1234}	0.125	0.47

We have ignored for the time being the fact that the temperature was held at the wrong level in trial 8 and the above regression coefficients are the same as those given in Table 2.C and Table 2.D. Those factors which are correlated with "blocks" all have relatively large regression coefficients associated with them. It will be noted that x_2 which is negatively correlated with "blocks" has a positive regression coefficient which suggests that it is an underestimate whereas the other larger coefficients are overestimated. The conclusion to be derived from this table is that the blocking effect is almost certainly the cause of the unexpected results obtained in the original analysis and the difference in the average response for trials 1-6 and trials 7-16 must be taken into account in the analysis.

As a preliminary analysis only the main effects were considered and the following coefficients were determined.

$$b_1 = -0.186$$
$$b_2 = 0.980$$
$$b_3 = 0.191$$
$$b_4 = 0.109$$
$$b_B = 4.098$$

The regression coefficients b_1, b_3 and b_4 are of equal magnitude and are appreciably smaller than b_2 and b_B. The latter refers to the "blocks". The significance of these effects is tested in the following analysis of variance table.

TABLE 2.F.

TESTING THE SIGNIFICANCE OF THE REGRESSION COEFFICIENTS

Source of Variation	D.F.	S.S.	M.S.
Blocks and Temperature	2	229.1	114.55
Extra due to Gear Pump, Kiln Speed and Primary Air	3	3.7	1.23
Regression	5	232.8	
Residual	10	20.6	2.06
Total	15	253.4	

The residual mean square, 2.06, is consistent with the value of 1.834 derived from the repeat readings for each trial. There is no evidence that the gear pump speed, the kiln speed or the primary air volume affect the grading and it now remains to see whether or not the temperature has a significant effect. Ignoring the other factors, the new estimates of the regression coefficients for temperature and "blocks" were determined to be 1.06 and 3.864 respectively.

TABLE 2.G.

ANALYSIS OF VARIANCE TESTING THE SIGNIFICANCE OF TEMPERATURE

Source of Variation	D.F.	S.S.	M.S.	F.
Blocks	1	212.3	212.3	114
Extra due to Temperature	1	16.8	16.8	9.0
Regression	2	229.1		
Residual	13	24.3	1.87	
Total	15	253.4		

The value of $F = 9.0$ on 1 and 13 degrees of freedom just reaches the 1 per cent level of significance and we conclude that the temperature does have a bearing on the grading. The "blocks" are of course highly significant. The average response for trials 1-6 was 72.89 compared with an average of 65.37 for the remaining trials, and this represents the outstanding effect as regards the grading. The regression coefficient for temperature was 1.06 which means that the average effect of increasing the temperature from 285° to 295° was to increase the response by 2.12 per cent. This increase is only about a quarter of the difference between the blocks, namely 7.72 per cent, and this suggests that though the temperature can be used to increase the grading there is a much greater potential gain if the exact causes of the changeover between trials 6 and 7 can be determined. As the residual mean square is consistent with the value of 1.83, it was concluded that the effects of the individual interactions must be negligible.

2.4. SECONDARY RESPONSE.

Although this experiment was primarily designed to investigate factors affecting the grading of Britmag there wase also a secondary response to consider. This was the bulk density of the product which has to be greater than 2.90 to meet the consumers' specifications. As stated earlier, seven samples of the product had been taken over a period of twelve hours, the first, fourth and last of these were analysed and the bulk modulus determined in each case. There were only two figures quoted for trial 1. From these results the standard error of a determination was calculated from the variance within the trials to be 0.0594. The standard error of a mean of three readings is thus $0.0594/\sqrt{3}=0.0343$; the variance being 0.00118.

The mean results obtained in each trial are tabulated below.

TABLE 2.H. * See facing page.

MEAN BULK DENSITIES

Trial	Average Bulk Density
1	2.88
2	2.85
3	2.84
4	2.94
5	2.92
6	2.89
7	2.74
8	2.82
9	2.75
10	2.91
11	2.74
12	2.76
13	2.69
14	2.69
15	2.70
16	2.75

Once again there is an obvious changeover between trials 1-6 and 7-16 and so the "blocks" were introduced into the analysis again. The regression coefficients for main effects and "block" were determined to be as follows:-

$$b_1 = -2.242 \times 10^{-2}$$

$$b_2 = -1.435 \times 10^{-2}$$

$$b_3 = 3.367 \times 10^{-2}$$

$$b_4 = -0.828 \times 10^{-2}$$

$$b_B = 6.219 \times 10^{-2}$$

The significance of these effects is tested in the following analysis of variance table.

TABLE 2.I.

ANALYSIS OF VARIANCE TESTING THE SIGNIFICANCE OF THE DIFFERENT FACTORS

Source of Variation	D.F.	S.S.	M.S.	F.
Blocks	1	0.06501	0.06501	61.3
Gear Pump Speed	1	0.00680	0.00680	6.42
Kiln Speed	1	0.01626	0.01626	15.3
Extra due to Temp. and Primary Air	2	0 .00416	0.00208	1.96
Regression	5	0.09223		
Residual	10	0.01057	0.00106	
Total	15	0.10280		

The residual mean square, 0.00106, is consistent with the value of 0.00118 derived from the variation within groups, which suggests that there is no need to examine the individual interactions as they will be of negligible account. Of the main effects there is no evidence to suggest that temperature or primary air affect the bulk density of the product, but the effect of gear pump speed is significant at the 5 per cent level while for kiln speed the significance reaches the 1 per cent level. Once again "blocks" are highly significant.

The regression coefficient for gear pump speed, b_1, is negative which implies that the lower level of speed gives the higher bulk density, but as the speed is measured as the number of seconds per 10 revolutions the lower level is in fact the higher speed. Increasing the speed over the range considered here increases the bulk density by 0.045 on the average. The effect of increasing the kiln speed is to increase the bulk density by 0.067. The average bulk density in trials 1-6 was 2.887 compared with 2.755 for the remainder, a difference of 0.132.

2.5. SUMMARY OF THE CONCLUSIONS

The effects of four factors, Gear Pump Speed, Back End Temperature, Kiln Speed and Primary Air Volume on the grading and bulk density of Britmag have been studied in this experiment. Of these factors only Back End Temperature was found to significantly affect the grading, increasing the temperature from 285°C to 295°C resulting in an increase of 2.12 per cent in the response. Temperature had no significant effect on the bulk density of the product but increasing the Kiln Speed from 0.59 revolutions per minute to 0.61 revolutions per minute caused the bulk density to increase by 0.067 on the average, while increasing the gear pump speed from 88 seconds/10 revolutions to 86 seconds/10 revolutions resulted in an increase of 0.047. There was no evidence that Primary Air Volume affected the grading or bulk density.

During the experiment, namely between trials 6 and 7, it had been observed that the initial paste had apparently become wetter and this coincided with a general decrease both in the grading response and the bulk density. The average for trials 1-6 inclusive and 7-16 were as follows:-

	Trials 1-6	Trials 7-16	Differences in Averages
Average Grading	72.89	65.37	7.52
Average Bulk Density	2.887	2.755	0.132

No further work was done on this problem due to this particular kiln having to be closed down for overhaul. A new air system was installed but it was decided in any case to carry out any further tests on No.4 kiln, which would allow a wider range of levels of the factors to be considered.

CHAPTER 3.

FACTORS AFFECTING BRITMAG GRADING - LABORATORY INVESTIGATION.

3.1. INTRODUCTION

It was proposed to study on a laboratory scale a process which was rather similar to that performed in the rotary kiln plant.

It is described in the introduction to Chapter 2 how a magnesia slurry is precipitated from a reaction between pretreated sea water and calcined dolomite and this precipitant, after filtration, is fed into the rotary kiln. As the filter paste passes through the kiln, which is heated by burning pulverised coal, water and alkali salts are volatilized and the residue fuses to form brownish granules which are known to the refractories industry of this country as Britmag.

The problem to be investigated is one of improving the grading of the Britmag, an improvement being a reduction in the amount of fine material in the final product.

It can be seen that there are roughly three stages in the kiln process. Firstly there is the drying stage where the free moisture is evaporated, secondly the calcination stage and finally the fusing, and it is principally on the first two stages that the ultimate size of the Britmag granules depends. This is because the drying and calcination determine the strength of the material at this stage and obviously the weaker the material at this stage the smaller the final granules of Britmag as the rotation of the kiln causes a disturbance of the material and consequently breakdown. It seems feasible therefore that an increase in the strength of the material after drying and calcination would result in an improvement in the grading of the Britmag and it was with this view in mind that the following experiment was carried out.

3.2. PILOT EXPERIMENT

The object of this investigation was to determine the effect of drying temperature, drying time, free moisture content, calcination and the influence of additives on the mechanical strength of the filter paste.

Some of the conditions imposed in this experiment were similar to those operating in the actual kiln process but were at different levels. For example, in this experiment the drying temperature ranged between 30°C and 140°C. whereas in the kiln the filter paste is dried by contact with gases at a temperature of about 300°C. Similarly the temperature at which the dry paste was calcined in the experiment was 500°C. compared with a temperature range of 400-1000°C. in the kiln. From different drying temperatures it was hoped to get some idea of the effect of having different rates of drying. It was known from previous experiments that calcination at 500°C. gave more friable material than calcination at any other temperature, and it was assumed that any improvement achieved at this temperature would hold at a higher level. This extrapolation is recognised as being a dangerous practice and not to be commended but as this is only a preliminary experiment it was hoped from a study of these effects to get an indication as to the type of experiments which would need to be carried out in future.

Two additives for which there was prima facie evidence that they would strengthen the paste were tested and to determine their effect paste with no addition was included as a control. The free moisture content was varied at two levels, which were actually measured by the MgO concentration of the paste. The paste was dried to the higher level of MgO concentration by pressing between plaster bats which absorbed some of the free moisture. The drying temperature was varied at four levels at each of which the drying time was varied over three levels, which were different for each temperature. These drying times were chosen so that the paste would be dry enough to be put to a standard milling test.

Having been subjected to a particular set of the previous conditions the dried paste was divided into two portions and by random selection one of them was given the standard milling test while the other was first calcined and then milled. This milling test was the means of assessing the mechanical strength of the paste and the procedure was to mill a constant weight of dried paste, having a standard size grading, with 250 grams of $\frac{3}{4}$ inch dia. steel balls in a six inch diameter steel cylinder, 8 inches in length, rotating at 20 revs./min. The proportion of material reduced to less than 1/16" in size was taken as a measure of the mechanical strength. A strong paste would thus have a small breakdown.

The investigation was in the nature of a pilot experiment and as it was not certain which factors would interact, it was decided to examine all combinations of the factors at their respective levels. Accordingly a complete replication of a $12 \times 3 \times 2^2$ factorial experiment, including 144 trials in all, was carried out with the intention of pooling second and higher interactions as an estimate of the residual variance. The cost of each trial was very small and there was consequently no need to carry out only a fractional replicate.

It was impossible to randomise this experiment completely and in order that it could be a practical proposition each block, where a block represents all treatment combinations for a particular drying temperature, was carried out in full and the order of the blocks randomised as well as the combinations within each block. In order to eliminate any possible bias in the filter paste, sufficient paste for a complete replication was obtained and having been sub-divided into 72 portions it was stored in heat-sealed polythene bags so that there should be no difference in the starting material for any treatment combination. Any differences which might occur would be confounded in the effect of drying temperature, but assurances were given that these differences should not arise. * See facing page.

The factors and their levels were as follows:-

Predrying (T)

Drying Temperature		Drying Time (hours) (1)	(2)	(3)
(1)	30°C.	18	24	30
(2)	60°C.	12	15	18
(3)	100°C.	8	10	12
(4)	140°C.	4	5	6

Treatment levels will be denoted by T_{ij} where the first suffix represents the level of drying temperatures and the second suffix the level of drying time. ($i = 1,2,3,4$; $j = 1,2,3$).

Free Moisture Content, (W)

(1) Untreated filter paste - 32/33 %MgO ; (Wo)

(2) Dried filter paste - 39/40 %MgO ; (W_1)

Additives, (A)

(1) Control ; (Ao)

(2) Sodium Silicate, 1% addition ; (A_1)

(3) Aluminium Orthophosphate, 1% addition ; (A_2)

Calcination, (C)

(1) No calcination ; (C_o)

(2) Calcined at 500°C. for 5 hours ; (C_1)

3.3. RESULTS AND ANALYSIS

The percentage breakdown of the dried filter paste in the standard milling test for each treatment combination is shown in Table 3.A.

TABLE 3.A.

PERCENTAGE BREAKDOWN IN MILLING TEST.

			T_{11}	T_{12}	T_{13}	T_{21}	T_{22}	T_{23}	T_{31}	T_{32}	T_{33}	T_{41}	T_{42}	T_{43}
C_o	W_o	a_0	59.8	37.8	36.8	39.3	38.3	46.3	31.3	49.0	47.3	62.0	71.8	49.3
		a_1	38.0	49.3	40.5	57.8	51.0	56.0	49.5	61.3	73.3	81.0	53.8	46.5
		a_2	48.8	42.3	39.8	45.0	51.5	53.8	50.3	57.0	60.3	72.8	83.0	25.8
	W_1	a_0	26.0	20.0	20.8	2.8	7.5	21.3	18.0	25.0	18.3	33.3	13.5	24.3
		a_1	21.0	16.0	25.0	7.5	41.3	21.8	22.5	15.8	29.0	3.8	22.0	25.0
		a_2	24.8	28.8	19.0	19.3	17.8	13.3	29.3	13.3	21.3	24.8	10.5	3.0
C_1	W_o	a_0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
		a_1	100.0	95.9	97.7	100.0	100.0	100.0	100.0	100.0	100.0	88.2	100.0	100.0
		a_2	100.0	98.6	98.5	94.9	96.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	W_1	a_0	63.6	42.4	94.4	94.2	57.9	68.8	51.2	98.0	73.5	54.3	42.5	86.3
		a_1	39.7	94.3	76.0	66.5	86.1	88.8	50.9	90.8	61.1	50.8	96.1	94.1
		a_2	67.0	87.4	95.6	99.2	98.4	59.6	77.9	97.2	80.0	93.8	79.8	88.1

At this point we were interested in the following questions:-

(1) What is the effect of drying the paste at different temperatures?

(2) What is the effect of varying the drying time at a particular temperature?

(3) Does the removal of free moisture have a beneficial effect on the strength of the paste?

(4) Do the additives have a beneficial effect and is this effect the same for both additives?

(5) What is the effect of calcining the paste?

(6) Are there any interactions between the treatments?

It can be seen from Table 3A that the paste at the lower MgO concentration, which had been calcined, had suffered 100 percent breakdown in nearly every case. The data in this block was therefore omitted from the analysis and the design modified to a 12 x 3^2 factorial experiment.

Denoting the treatment combinations C_oW_o, C_oW_1, C_1W_1 by B_o,B_1,B_2, respectively, the analysis of variance testing the general treatment effects as shown in Table 3.B.

TABLE 3.B.

ANALYSIS OF VARIANCE OF TABLE 3.A.

Source of Variation		DF	SS	MS	F
Main	T	11	1,602.56	145.69	1.14
Effects	A	2	932.29	466.15	3.64
	B	2	58,132.50	29,066.25	227.04
Interactions	TA	22	4,095.98	186.18	1.45
	TB	22	7,658.09	348.10	2.72
	AB	4	1,103.93	275.98	2.16
Residual (Higher order interactions)		44	5,632.72	128.02 (= σ 2)	
Total		107	79,158.01		

In the absence of replication or an independent estimate, the higher order interactions mean square is used as an estimate of the error variance. The analysis shows that B is very highly significant and A is significant at the 5 percent level, while the interaction TB, is also significant at the 1 percent level. None of the other effects reach the 10 percent level of significance. * See facing page.

Neither of the interactions TA or AB are statistically significant. The effects of the two additives can therefore be determined by averaging the results over all levels of T and B. This gives the following table of mean percentage breakdown, each entry being the average of 36 results.

TABLE 3C.

EFFECT OF THE TWO ADDITIVES (MEAN % BREAKDOWN)

Control	Sodium Silicate	Aluminium Orthophosphate
45.18	50.07	52.19

The standard error of the difference between any two of these means is 2.67. There is therefore no significant difference in the effect of the two additives, and the use of either of these additives results in a significantly higher percentage breakdown than is obtained for untreated paste.

In analysing the effects of the predrying treatments, the removal of free moisture and the calcination, it is helpful to examine the following table. The entries in this table are the mean percentages breakdown for paste with and without additives.

The two major effects are due to the initial removal of some free moisture and to the calcination. The average effect of increasing the MgO content was to reduce the breakdown from 51.6% to 19.6% for uncalcined paste, while for calcined material the corresponding reduction was from 99.2% to 76.3%. Calcining the paste under any conditions caused a large increase in the amount of breakdown.

Drying Temp. (°C)	Drying Time(hrs)	Not calcined		Calcined	
		32/33% MgO	39/40% MgO	32/33% MgO	39/40% MgO
30	18	48.9	23.9	100.0	56.8
	24	43.1	21.6	98.2	74.7
	30	39.0	21.6	98.7	88.7
60	12	47.0	9.8	98.3	86.6
	15	46.9	22.2	98.9	80.8
	18	52.0	18.8	100.0	72.4
100	8	43.7	23.3	100.0	60.0
	10	55.7	18.0	100.0	95.3
	12	60.3	22.8	100.0	71.5
140	4	71.9	20.6	96.1	66.3
	5	69.5	15.3	100.0	72.8
	6	40.5	17.4	100.0	89.5
Average Breakdown		51.6%	19.6%	99.2%	76.3%

The magnitude of the effects of removing free moisture and calcining the paste varied for the different predrying treatments, the interaction TB being significant. The effect of varying the drying time was therefore examined for each of the drying temperatures separately, ignoring the results for calcined paste with an initial content of 32/33% MgO. For paste dried to 39/40% MgO, and not calcined, there was no significanteffect of drying time or temperature and all the results obtained under these conditions were comparable. For the remaining pastes the effect of the drying time varied at the different drying temperatures. In some cases increasing the drying time increased the percentage breakdown, while in other cases increasing the drying time reduced the breakdown. It would appear therefore that under given conditions there is a particular drying time for which the resulting paste is weakest. This is an interesting result.

3.4. DETERMINATION OF OPTIMUM CONDITIONS FOR PREDRYING PROCESS AND MECHANICAL WORKING.

One of the most important indications to come from the pilot experiment was that the removal of free moisture by pressing between plaster bats increased the strength of the filter paste. The removal of the free water by this method was not standardised in the last experiment and the question arose whether or not the actual working of the paste, which was a necessary part of the pressing with the plaster bats, was also a contributory factor. It was decided to investigate this and a few preliminary trials were carried out which indicated that working the paste after it had been partially dried did have a beneficial effect.

As it was hoped at some future date to put the results of these experiments to the test on a plant scale, the previous method of free moisture removal by pressing between plaster bats was discarded and the water removed by drying in an oven at a fixed temperature. It was proposed to dry the paste to a certain concentration of MgO, work it mechanically, and then dry it again until the total drying time was six hours, before finally calcining it.

It would be advisable at this stage to give an outline of the practical side of this experiment in order to bring to light the sources of error.

A sample of paste is taken from the current stock on the plant and 1,000 grams of this is taken and spread on a tray to a thickness of approximately 3/4 inch. The tray is then put in an oven and dried at 140°C. for half an hour, say, in which time the MgO concentration of the paste would increase from about 33% to 40%. The paste would then be worked mechanically by passing it through a mincing machine a required number of times and then spread on the tray again to a thickness of approximately $\frac{3}{4}$ inch and put back in the oven for $5\frac{1}{2}$ hours, thus making the total drying time 6 hours. The dried paste is then calcined at 500°C. for 5 hours before being subjected to a standard milling test, which is the means of assessing the mechanical strength of the final paste.

The first source of error arises from the fact that when the paste is put in the oven to dry, or calcined for that matter, it is not spread to exactly the same thickness everytime and so the drying effect is never really the same for each portion of paste. This has the added disadvantage that one cannot tell beforehand with any degree of accuracy what the MgO concentration is going to be before mechanically working it. A second source of error is due to the possible variations in the original filter paste which, as stated earlier, is taken from the current stock when required. It would have been possible to have obtained enough paste at one time, supposedly of homogeneous consistency, and stored it in polythene bags until it was required for use, but that would have meant the results have applied only to paste with the same physical and chemical properties as the one used. By taking paste from the plant stock when it was required, it meant the results could safely be assumed to hold in general for any paste, and this advantage more than offset the increase in the experimental variation.

A further source of error is due to the effect of the mechanical working not being exactly measurable by the number of times the paste is passed through the mincer. It was noticed after experimentation had begun that when the paste had been dried to a high concentration of MgO it was necessary to force the paste through the mincer manually, so that the amount of work done on the paste did not increase linearly with the number of times the paste was passed through the mincing machine.

These additional sources of error, which inflated the inherent experimental error, meant that more trials would need to be carried out than would normally have been required, but experimentation was not so costly as to make this a major point of consideration.

The predrying treatment was usually applied to four samples of paste at the same time; they were then removed and by random selection given a stipulated number of mechanical workings before being returned to the oven until the total drying time for each was six hours. They were then all calcined together. This meant that four trials could be carried out in little more than the time taken to do one trial separately.

3.5. METHOD OF ANALYSIS AND EXPERIMENTAL RESULTS

The title of this particular work (see 3.4) is a trifle misleading as we were not really interested in finding an optimum point, at which the paste is strongest for a particular combination of MgO concentration and the number of mechanical workings, but rather in studying an "optimum area". However, the general principles of Box's technique, as described in Chapter 1, do carry over to a certain extent as it was proposed to carry out a small initial group of experiments which would indicate in which direction the optimum area lay, and then move into it. Having got there it was intended to study this area initially by estimating an equation of the second degree and reducing it to its canonical form in order to assess visually the effects of pre-drying and mechanically working the paste.

The general principles of the sequential method of procedure will therefore be maintained but it is impossible in this experiment to make use of the principles of factorial design which are so essential for the simplicity of computation and analysis and in its place we will have to employ the methods of multiple regression at length. This is not so inconvenient as it may first appear as we are only dealing with two variables and an equation of the second degree would only require five regression coefficients to be estimated, which is not too troublesome on a desk calculator. Estimation of a surface of the third degree would be more of a problem but it was thought that this situation would not arise.

In the first experiment four trials were carried out in which two samples of paste were pre-dried for $\frac{1}{4}$ hour and two for $\frac{1}{2}$ hour before being given a set number of mechanical workings, the whole being replicated to overcome the large error which is known to exist. As stated earlier it was impossible to gauge the MgO concentration from the length of time it was pre-dried and it can be seen from the results in Table 3.E. that the second "replication" was at a much higher MgO concentration than the first. This is not disadvantageous as the method of analysis is that of multiple regression, though there was the possibility that too great a range of MgO concentration had been covered and would consequently necessitate the inclusion of second order terms immediately.

TABLE 3.E.

FIRST EXPERIMENT

Trial	% MgO x_1	No. of Mechanical workings x_2	% Greater than 1/16" after Milling. y
1	36.24	1	30.9
2	37.43	4	46.4
3	42.75	1	42.2
4	40.95	4	43.0
5	39.06	1	32.8
6	38.65	4	43.2
7	44.03	1	57.4
8	44.03	4	59.0

Assuming that first order effects are predominant an equation of the form

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

was set up as a model and this gave the equation:-

$$y = 2.63x_1 + 2.58x_2 - 68.3 \qquad \ldots\ldots\ldots\ldots(1)$$

The analysis of variance testing the significance of this regression plane as well as showing the effect of including both factors, is shown in Table 3.F.

TABLE 3.F.

ANALYSIS OF VARIANCE TESTING MULTIPLE REGRESSION.

Source of Variation	DF	SS	MS	F
Regression (x_1 alone	1	416.42	416.42	11.91 *
extra due to x_2	1	119.72	119.72	3.42
Regression for x_1 & x_2	2	536.14	268.07	7.67 *
Residual	5	174.84	34.97	
Total	7	710.98		

* Significant: $1\% < P < 5\%$

It can be seen from the above analysis that the hypothesis that first order effects are predominant at this stage is borne out by the significance of the regression. The analysis also shows that x_1 is the dominant factor and that the inclusion of x_2 does not significantly improve the fit of the regression plane.

The indication at this stage therefore is to increase the MgO concentration of the paste before working it. The effect of the mechanical treatment is not so obvious and this may be because

(a) it has no effect,

(b) there is not enough difference between 1 and 4 workings to show a significant effect,

(c) the base level chosen for this particular factor is near a conditional maximum.

It was known that the paste must be given some mechanical working, unworked paste giving 100 per cent breakdown in all cases, but it was not known at this stage whether or not a minimum amount of working is required after which it has no effect on the strength of the paste. It was proposed to accept temporarily that an increase in the number of mechanical workings did strengthen the paste and following the line of steepest ascent, further trials were carried out to see whether the indications of the first eight trials were confirmed. Accordingly it was suggested that the paste be dried for 40 minutes and worked six times and dried for 45 minutes and worked seven times, each of the trials replicated. Thes results obtained were as shown in Table 3.G., in trials 9-12.

TABLE 3.G.

PATH OF STEEPEST ASCENT AND SECOND EXPERIMENT

Trial	%MgO x_1	No. of Mechanical Workings x_2	Response y
9	47.15	6	78.3
10	47.15	6	69.8
11	45.25	7	63.5
12	47.24	7	70.8
13	50 (Approx)	8	-
14	50 (Approx)	8	Paste too dry for Mechanical Treatment
15	42.72	2	60.6
16	43.97	4	34.0
17	41.90	6	34.0
18	42.82	8	64.6
19	43.39	2	15.4
20	46.24	4	83.1
21	43.86	6	72.4
22	45.71	8	76.9

A marked increase in the strength of the paste is obvious from a visual inspection of the results. Continuing along the line of steepest ascent trials 13 and 14 were found to give a paste so dry as to be unable to work it - the dried paste having crumbled to a powder. This sort of thing was not unexpected and it was decided to move back from this critical region to a safer area. The area elected upon was one for which the paste concentration ranged between 43 and 47 percent MgO and to try to establish just what the effect of mechanical working was, eight further trials were performed in which the paste was dried for 40 minutes and 45 minutes and for each drying time four samples were worked 2,4,6 and 8 times respectively. Three of these trials fell outside the area in which we were interested and four further trials were included as supplementary points.

TABLE 3.H.

SUPPLEMENTARY POINTS

Trial	% MgO x_1	No. of Mechanical Workings x_2	Response y
23	45.41	2	87.8
24	44.88	4	72.5
25	46.43	6	81.9
26	46.03	7	55.2

The area to be investigated was suspected of being large enough to necessitate the inclusion of second order effects. Accordingly, all points lying in this area (and there are 15 in all, taking into consideration those points in the earlier trials), were used to estimate an equation of the form:-

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2$$

and this gave rise to the equation:

$$y = 45.87x_1 - 1.39x_2 - 8.21\, x_1x_2 + 1.29x_1^2 + 1.88x_2^2 - 1901.34 \quad \ldots (2)$$

The analysis of variance testing the goodness of fit of equation (2) is shown in Table 3.I.

TABLE 3.I.

ANALYSIS OF VARIANCE FOR MULTIPLE REGRESSION.

Source of Variation	DF	SS	MS	F
Multiple Regression	5	4168.28	833.66	7.05
Residual	9	1063.58	118.18	
Total	14	5231.86		

The variance ratio testing the adequacy of equation (2) as a representation of the data is 7.05 on 5 and 9 d.f. which is significant at the 1 percent level. It is of interest at this stage to test (a) if second order terms are important, (b) if the effect of mechanical working is significant. This is done by partitioning the s.s. due to regression as follows:-

Testing the need to include second order terms:

Source of Variation	DF	SS	MS	F
Linear Effects	2	2005.50	1002.75	8.45 **
Extra due to 2nd order terms	3	2162.78	720.93	6.10 *
Multiple Regression	5	4168.28		
Residual	9	1063.58	118.18	
Total	14	5231.86		

Testing the need to include mechanical working:

Source of Variation	DF	SS	MS	F
Ascribable to x_1 & x_1^2	2	3052.36	1526.18	12.91 **
Extra due to terms with x_2	3	1115.92	371.97	3.15 Ψ
Multiple Regression	5	4168.28		
Residual	9	1063.58	118.18	
Total	14	5231.86		

** Highly significant : $P < 1\%$

* Significant : $1\% < P \leqslant 5\%$

Ψ Approaching ~~Possibly~~ significante: $5\% < P \leqslant 10\%$

It can be seen from the above two tables that the inclusion of second order terms significantly improves the adequacy of the regression equation, while the importance of the mechanical working is still not definitely established, though it does appear likely that there is an effect.

The above calculations show that an equation of the second degree is adequate to describe the response surface in the region of the design but in its present form equation (2) does not convey a great deal about the nature of the response surface. It is proposed to plot a contour diagram for the yields and to facilitate the procedure, the equation is first reduced to its canonical form

$$Y - Ys = B_{11}X_1^{\,2} + B_{22}X_2^{\,2},$$

where Ys is the response at the centre of the system, S, which is taken as the new origin and X_1 and X_2 are the principle axes of the new system.

The centre, S, is found by differentiating equation (2) with respect to x_1 and x_2 and equating to zero. This gave the co-ordinates of the centre as

$$x_{1s} = 45.8 \qquad\qquad x_{2s} = 6.5$$

with a corresponding yield Ys = 68.01.

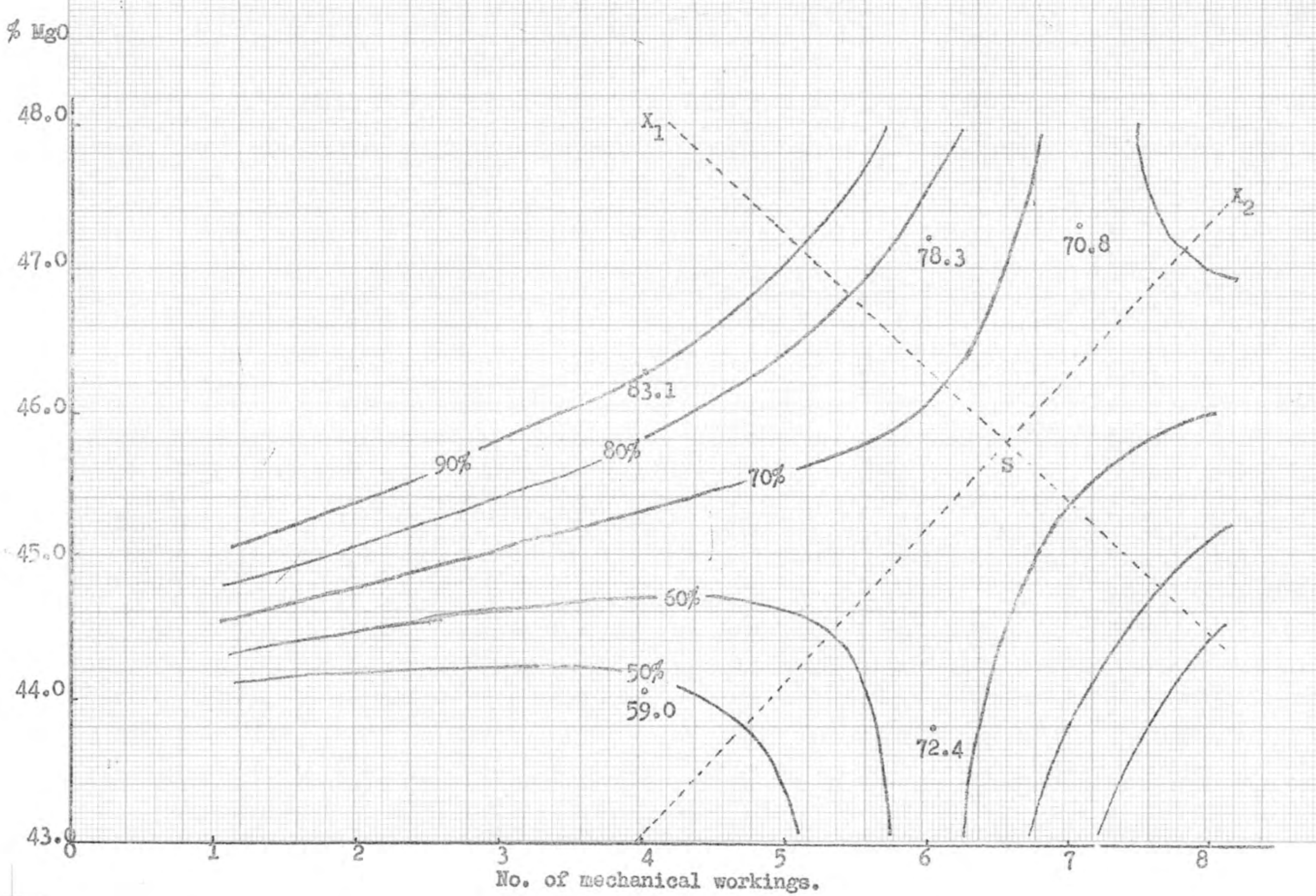
The direction of the axes, X_1 and X_2, and the values of the constants B_{11} and B_{22} were determined in the usual way.
This gave the equation

$$Y = 5.70\,X_1^{\,2} - 2.53\,X_2^{\,2} + 68.01 \qquad \ldots\ldots\ldots\ldots\ (3)$$

The geometrical representation of equation (3) is illustrated in Figure 3.1., and we can see that we have a saddle-point (sometimes called a minimax or col), which is what was expected as the coefficients B_{11} and B_{22} are of the same order of magnitude and with opposite signs.

FIGURE 3.1.

Contours of fitted surface and some observed yields at experimental points.



The important indications for the contour diagram at this stage are as follows:-

(a) Moving along the X_1 axis away from the centre gives a higher response, which means we can either have a dry paste with a few number of workings or a wetter paste with more workings.

(b) A dry paste with a large number of workings will not give a strong paste.

However, we are really extrapolating outside the experimental region as the centre of the new co-ordinate system lies towards the outside of the experimental design and it was therefore decided to carry out further trials to supplement the present design in order to confirm or disprove the present indications. It was therefore recommended that additional trials be performed in the regions of particular interest, namely for paste dried to 43-44% MgO with 7-9 mechanical workings, and for paste dried to 45-46% MgO with 2 and 3 workings. Several trials were carried out as suggested but due to the experimental difficulties some of the trials did not fall in the areas intended, even though they did fall in the general region of interest. Nothing definite could be established from a visual study of these new results, and these additional observations were therefore pooled with the earlier data, and a new response equation of the second degree was estimated to see whether the suggested tendencies were still in evidence.

Recalculation of the equation gave:-

$$y = 34.66x_1 + 9.30x_2 - 2.47\,x_1x_2 - 3.15\,x_1^2 - 0.23x_2^2 + 5390.5 \quad \ldots . \,(4)$$

The analysis of variance testing the adequacy of this equation as a representation of the data is shown in Table 3.J., together with the effect of including mechanical treatments.

TABLE 3.J.

ANALYSIS OF VARIANCE TESTING MULTIPLE REGRESSION

Source of Variation	DF	SS	MS	F
Ascribable to x_1 and x_1^2	2	2709.771	1354.886	8.45
Extra due to terms with x_2	3	1879.023	626.341	3.91
Multiple Regression	5	4588.794	917.759	5.73
Residual	31	4968.204	160.265	
Total	36	9556.997		

For multiple regression, the value of F is 5.73 on 5 and 31 degrees of freedom which is significant at the 1 percent level, while the variance ratio testing the effect of including terms dependent upon x_2 is 3.91 on 3 and 31 degrees of freedom which just reaches the 2 percent level of significance. Thus the mechanical working does have a definite effect on the final strength of the paste.

Once again the regression equation was reduced to its canonical form and this gave:-

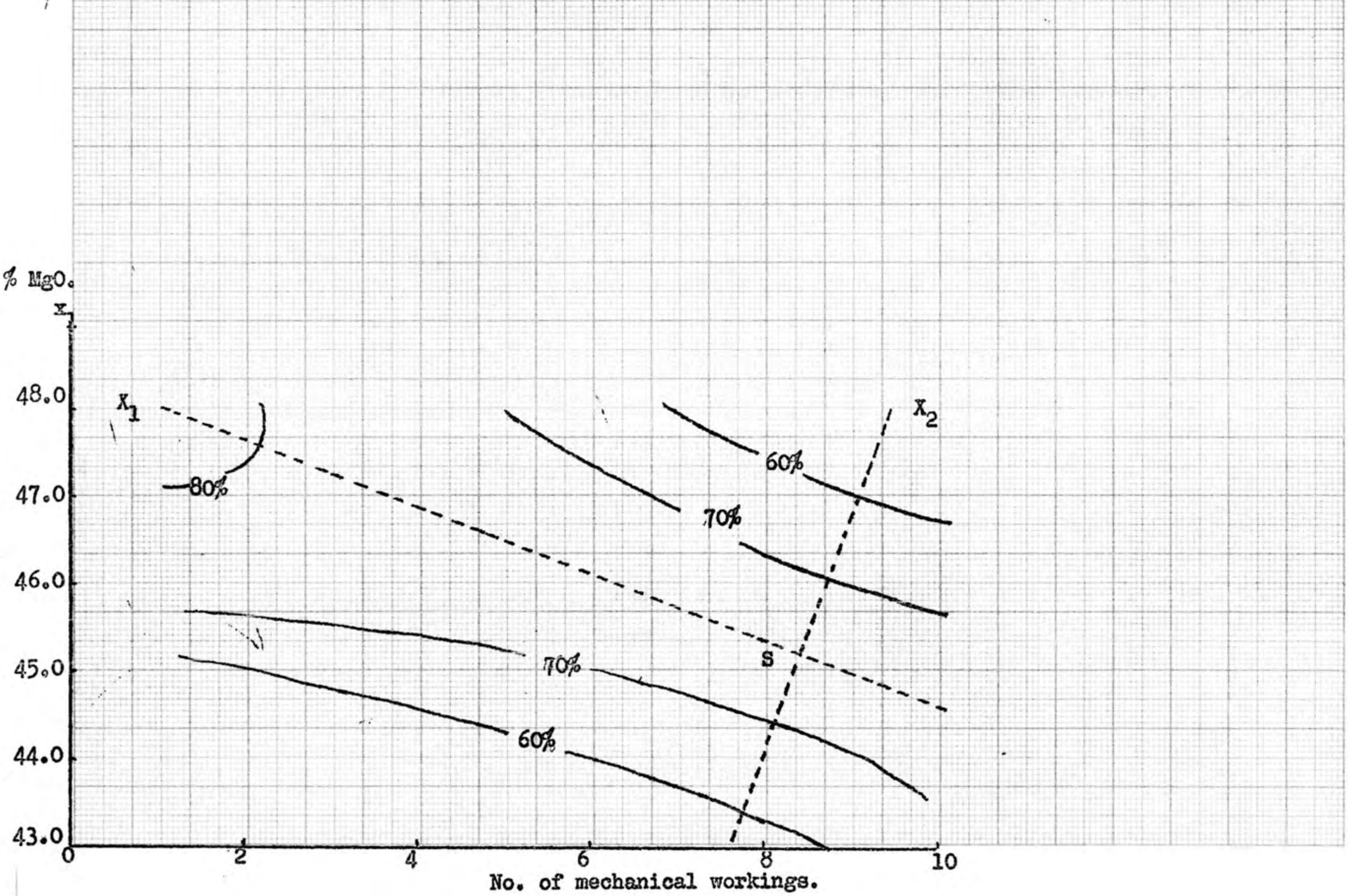
$$y = 0.225\ X_1^2 - 3.597\ X_2^2 + 72.95$$

with the new centre, S, at $x_1 = 45.24$ and $x_2 = 8.34$

For diagrammatic purposes the response contours were replotted and this gave the situation as illustrated in Figure 3.2. There is an obvious change from the earlier representation, the main characteristics being

FIGURE 3.2.

CONTOURS OF REFITTED SURFACE.



emphasised by the relative magnitudes of the coefficients in the canonical form. The centre of the co-ordinate system is still some distance from the centre of the experimental region and the situation is becoming more like that of a stationary ridge.

B_{22} is appreciably larger than B_{11} and a good approximation to the surface can be obtained by taking B_{11} to be zero. The equation of the surface then becomes

$$y = 72.95 - 3.6\, X_2^{\,2}$$

$$\text{or} \qquad X_2 = \pm \sqrt{\frac{72.95 - y}{3.6}} = \pm\, 0.527 \sqrt{72.95 - y}$$

This equation defines a pair of parallel planes on which the response is y. This would indicate a plane $X_2 = 0$ which is maximal for which the response is 72.95.

A review of the practical implications is warranted at this stage in order to explain the logic of the trials already performed and those to be carried out. It has been established that pre-drying the paste and working it mechanically before calcination, improves the strength of the paste. We are now interested in finding a combination of the two factors for which the strength of the paste is satisfactory without necessarily being maximal. This last statement is not as ludicrous as it may at first appear, for in performing this sort of process on a plant scale it would be much easier to work the paste than to pre-dry it to the range of MgO concentration we are dealing with here. If a dry paste with a small amount of working gave the same, or only slightly better, results than a wetter paste with more workings, then the latter conditions would be preferable from a practical point of view. Thus it is more important at this stage to determine whether the indication that more workings for relatively wetter paste give good results is borne out away from the area already under consideration.

With this is mind, a further set of trials were performed in which the MgO concentrations were decreased down to approximately 39 percent with corresponding increases in the number of mechanical workings. These are tabulated in Table 3.K. though not necessarily in the order in which they were carried out.

TABLE 3.K.

SUPPLEMENTARY TRIALS

% MgO x_1	No. of Mechanical Workings x_2	Response y
43.00	10	56.8
42.99	10	75.4
42.20	10	52.8
41.72	13	45.9
41.66	15	63.6
41.59	15	62.3
41.39	13	64.8
41.28	11	62.8
41.23	13	71.2
40.83	11	74.6
40.40	15	61.7
39.09	20	55.1
39.23	25	59.8
38.91	20	53.3
39.11	25	65.6

These results, although very variable, confirm the tendency indicated that wetter paste with a larger amount of working is comparable in strength with a paste dried to a higher level of MgO concentration with relatively less working. This was a very favourable state of affairs for the reasons already mentioned, and few further tests were performed on undried paste, of approximately 34% MgO, with 20 and 25 mechanical workings. These experimental conditions gave yields in the region of 20%, which, although not as good as those obtained for drier paste, are of some interest. The number of mechanical workings for paste concentrations of 34 percent can hardly be related to those for drier paste as the relative amount of work done is so different. In order to investigate an area in which wet paste is used, it will be necessary to find a new method of working the paste which would allow more accurate comparisons to be made. It was decided therefore, that this investigation had served its purpose in giving an insight into the system and so experimentation was adjourned.

As opposed to predrying the paste and then working it mechanically in order to increase the strength of the final product, a series of experiments was carried out in which the paste was treated with additives to see if comparable results could be obtained. In these trials ten chemicals were added at two levels to ordinary filter paste, and the strength of the dried and calcined paste measured by the percentage breakdown to less than 1/16 inch. In every case one hundred percent breakdown was recorded, and it was therefore concluded that the additives were less efficient than the predrying and working method in strengthening the paste.

GENERAL CONCLUSIONS AND RECOMMENDATIONS.

From the investigation described above, it was determined that predrying the filter paste and working it mechanically affected the strength of the final product. The results indicate that predrying the paste before working it is very beneficial, but that comparable results may be obtained from a relatively wetter paste subjected to more intensive working. The experimental technique and the differences in the properties of the initial filter paste meant that the results obtained were necessarily very variable, but nevertheless, it was possible to draw some reliable conclusions. For material which had been dried upto 46% MgO and worked twice or thrice, the average yield (measured by the percentage of the final paste not broken down to less than 1/16 inch in size in the standard milling test), was about 70% and for paste with an MgO concentration of 39% MgO, which had been worked approximately twenty times, the average response was about 60 percent.

However, these results in themselves are hardly strong enough to enable one to say it would be worthwhile to employ a similar set-up on a plant scale, and this should be regarded purely as a preliminary enquiry. It may be argued that the simplicity of the conclusions did not warrant the need of such a long experiment with the corresponding excessive statistical analysis. However, before this investigation was carried out, it had been contended that only a certain number of mechanical workings would be required after which any increase in working would have no additional effect. This is probably true but it was not known what this number of workings is for any particular paste. In the preliminary experimentation, paste which had not predried and had received only half a dozen workings, resulted in 100 percent breakdown. It had then been concluded, wrongly, that the paste must be predried to a certain level before any improvement could be obtained. Therefore if this technique had not been employed, it may well have been that the results for paste of low MgO concentration with a higher number of mechanical workings would never have materialised, and consequently a very useful result overlooked.

It would be profitable to extend the above analysis using a different experimental arrangement, in which the effect of mechanical working is studied in further detail at much higher levels. A mechanical stirrer or milling operation would probably be much less susceptible to error than the method used previously as the operator effect will be excluded.

It is difficult to assess what a decrease in the amount of breakdown in the above tests is equivalent to on a plant scale. It certainly means that the paste is stronger, but what is this increase in strength equivalent to in terms of improvement in Britmag grading? The economicsof this problem must be carefully studied, and it should be ascertained quite definitely that the cost of reorganising the present system will pay dividends in a higher quality product.

CHAPTER 4.

INVESTIGATION OF DUST LOSSES FROM A ROTARY KILN AT COXHOE

4.1. INTRODUCTION

One of the basic raw materials used in the production of Britmag is Dolomite, a magnesium limestone, which is quarried at Coxhoe, a few miles from Hartlepool. This stone which contains approximately equimolecular amounts of magnesium carbonate and calcium carbonate is first crushed and screened before being calcined in a rotary kiln to give a mixture of magnesium and calcium oxides; the actual chemical process being

$$Mg.\ CO_3.\ CaCO_3 \rightarrow MgO.\ CaO + 2CO_2.$$

Following the crushing and primary screening operations, the stone fraction less than $1\frac{1}{2}$ inches in size is classified on a double-deck Nordberg screen fitted with $\frac{3}{4}$ inch and $\frac{3}{8}$ inch meshes. The material retained on the $\frac{3}{4}$ inch mesh and having a nominal size grading $+\frac{3}{4}'' - 1\frac{1}{2}''$ is conveyed directly to a storage silo of approximately 800 tons capacity. The fraction retained on the $\frac{3}{8}$ inch mesh is fed to a wash plant where material less than $\frac{1}{4}$ inch in size is removed in suspension and the residue, having a nominal size grading $+\frac{1}{4}'' - \frac{3}{4}''$, is conveyed to a second storage silo. Stone is drawn from either of these silos by Pendan weight feeders and conveyed to the rotary kiln. The unwashed stone used is equivalent to approximately 80% of the total consumption and is therefore used for much longer periods than the washed stone. The kiln is fired with pulverised coal and the stone decomposes during passage through the kiln, the magnesium and calcium carbonates being converted to the oxides with the evolution of carbon dioxide. The calcined product from the kiln is called Dolime.

The gases arising from the combustion of the coal and the decomposition of the dolomite pick up any fine particles of material and carry them out of the kiln. These gases are cleaned before being exhausted to the atmosphere by passage through cyclones which separate the dust from the gas stream; this dust loss from the kiln is waste material. It must be explained here that the company pays for the stone according to the weight of stone which is fed into the kiln, and therefore all dust loss from the kiln is a direct financial loss to the company. The dust loss would be expected to increase as the proportion of fine material present in the stone fed to the kiln increases. The purpose of the initial screening is to remove dirt and stone dust from the dolomite before it reaches the silo. However, the screening is not fully efficient and some fine material is carried over to the silo where further breakdown in the stone occurs due to the crushing effect of the weight of stone in the silo. Breakdown also occurs in the kiln due to the continual movement of the stone. In both cases the amount of breakdown is dependent upon hardness of the stone; so that stone hardness is a likely factor.

A practical point in the operation of the kiln is the formation of "clinker rings". A ring caused by coal ash and dolime fusing together, forms on the inside of the kiln and gradually builds up until it severely restricts the flow of material along the kiln. When this happens the kiln has to be stopped and the ash ring is removed by shooting it off with a gun. This happens about once every four days.

A series of experiments on dolime production, designed and analysed by N. Heasman, was carried out at Coxhoe in January. One of the problems investigated was the amount of dust loss from the rotary kiln when the unwashed stone was being used, and the object of the experiment was to establish a mathematical model to determine which factors affect the amount of dust loss. Previous tests on the dust plant effluent had shown the dust was mainly uncalcined dolomite having a particle size less than 100 B.S. mesh i.e. 0.006". It was therefore expected that the weight of dust loss would be directly related to the amount of dust in the feed to the kiln, and to assess the latter samples were taken hourly and the fraction less than 100 B.S. mesh was noted. The weight of the stone dust in feed, or stone "fines" entering the kiln per minute was then calculated from the mean Pendan feeder setting.

As stated earlier it was suspected that the stone hardness affected the dust loss and this was measured by percentage breakdown of material to a size less than 1/16 inch, in a standard milling procedure. Accordingly Heasman carried out determinations of hardness on the hourly samples from the Pendan feeder and tabulated mean values for each shift of eight hours duration. The amount of dust discharged from the cyclones was measured at hourly intervals by collecting and weighing the dust over a one-minute test period, and an average value per shift recorded.

Table 4.A. gives the observations on the dust loss (lbs./min.), Y, stone fines in feed (lbs./min.), X, and stone hardness, T, over the test period with the corresponding dates. Shifts when washed stone was being used have been omitted. The shutting down of the kiln for shooting off the clinker ring is also indicated.

Heasman estimated a multiple linear relationship of the type:

$$(y - \bar{y}) = \beta_1(x - \bar{x}) + \beta_2(t - \bar{t})$$

and this gave the equation:-

$$y = 0.5701x + 0.9642t + 47.4 \qquad (1)$$

The significance of this regression is shown in Table 4.B.

TABLE 4.A.

OBSERVATIONS

Date & Time		Stone Fines (lbs./min.) X	Stone Hardness (% breakdown) T	Dust Loss (lbs./min.) Y
30.1.57	6 - 2	25	17	109
	2- 10	43	17	108
	10 - 6	50	17	119
31.1.57	6 - 2	37	14	96
	2 - 10	43	14	100
1.2.57	6 - 2	36	14	96
	10 - 6	61	19	116
2.2.57	6 - 2	49	19	104
			CLINKER RING	
3.2.57	6 - 2	33	17	95
	2 - 10	73	17	109
	10 - 6	58	19	79
4.2.57	6 - 2	31	15	92
	2 - 10	43	15	86
	10 - 6	64	15	90
5.2.57	6 - 2	42	14	70
	2 - 10	34	14	71
	10 - 6	27	14	60
6.2.57	6 - 2	18	14	64
			CLINKER RING	
7.2.57	6 - 2	25	15	59
	2 - 10	33	15	69
	10 - 6	30	16	65
8.2.57	6 - 2	39	19	67
	2 - 10	38	16	74
			CLINKER RING	
10.2.57	10 - 6	56	22	70
11.2.57	6 - 2	45	29	94
	10 - 6	76	28	132
12.2.57	6 - 2	50	38	127
	2 - 10	56	37	105
	10 - 6	47	36	119
13.2.57	6 - 2	49	32	116
	2 - 10	54	32	109
	10 - 6	66	34	93

TABLE 4.B.

ANALYSIS OF VARIANCE FOR MULTIPLE REGRESSION

VARIATION	D.F.	S.S.	M.S.
Due to x and t	2	5,747.71	2,873.86
Unaccountable	29	7,944.01	273.93
Total	31	13,691.72	

The variance-ratio testing the adequacy of the fit is significant at the 0.1 percent level. To test whether each variable contributes significantly in the regression the analysis was extended as shown in table 4.C.

TABLE 4.C.

ANALYSIS OF VARIANCE FOR MULTIPLE REGRESSION

VARIATION	D.F.	S.S.	M.S.
Due to t	1	4,182.56	4,182.56
Extra due to x	1	1,565.15	1,565.15
Due to x and t	2	5,747.71	2,873.86
Unaccountable	29	7,944.01	273.93
Total	31	13,691.72	

The variance-ratio testing the effect of including x is thus significant at the 2 percent level. Thus x contributes a significant amount to the overall regression. It can be similarly shown that t contributes a significant amount to the regression. This does in fact verify Heasman's initial assumptions as the indications are that an increase in stone fines in the feed to the kiln will give rise to an increase in dust loss from the kiln, and an increase in t, which is the percentage breakdown to a certain size in a standard milling test and therefore a measure of stone softness rather than stone hardness, will also cause an increase in dust loss from the kiln.

Heasman then plotted the observed values of the dust loss and also the values predicted from equation (1). This is shown in Figure 4.1 This graph showed up certain irregularities and it is the explanation of these which is to be described in the remainder of this chapter.

4.2. EXTENSION OF HEASMAN'S ANALYSIS

It can be seen that the data falls into four groups, the boundary mark between successive groups being the stoppage of the kiln to shoot off the ash-ring. The two most striking features of Heasman's graph are the pronounced similarity in trends and the displacement of the observed and predicted values of the data in Groups I and III. These two features rather suggest that the individual regression planes for these two groups are probably parallel but not coincident. Groups II and IV are rather harder to interpret visually as they show much more variability though the later observations in Groups II do show a systematic difference between observed and predicted values. It was decided therefore to compare the regressions of the four groups to see if they do in fact differ significantly.

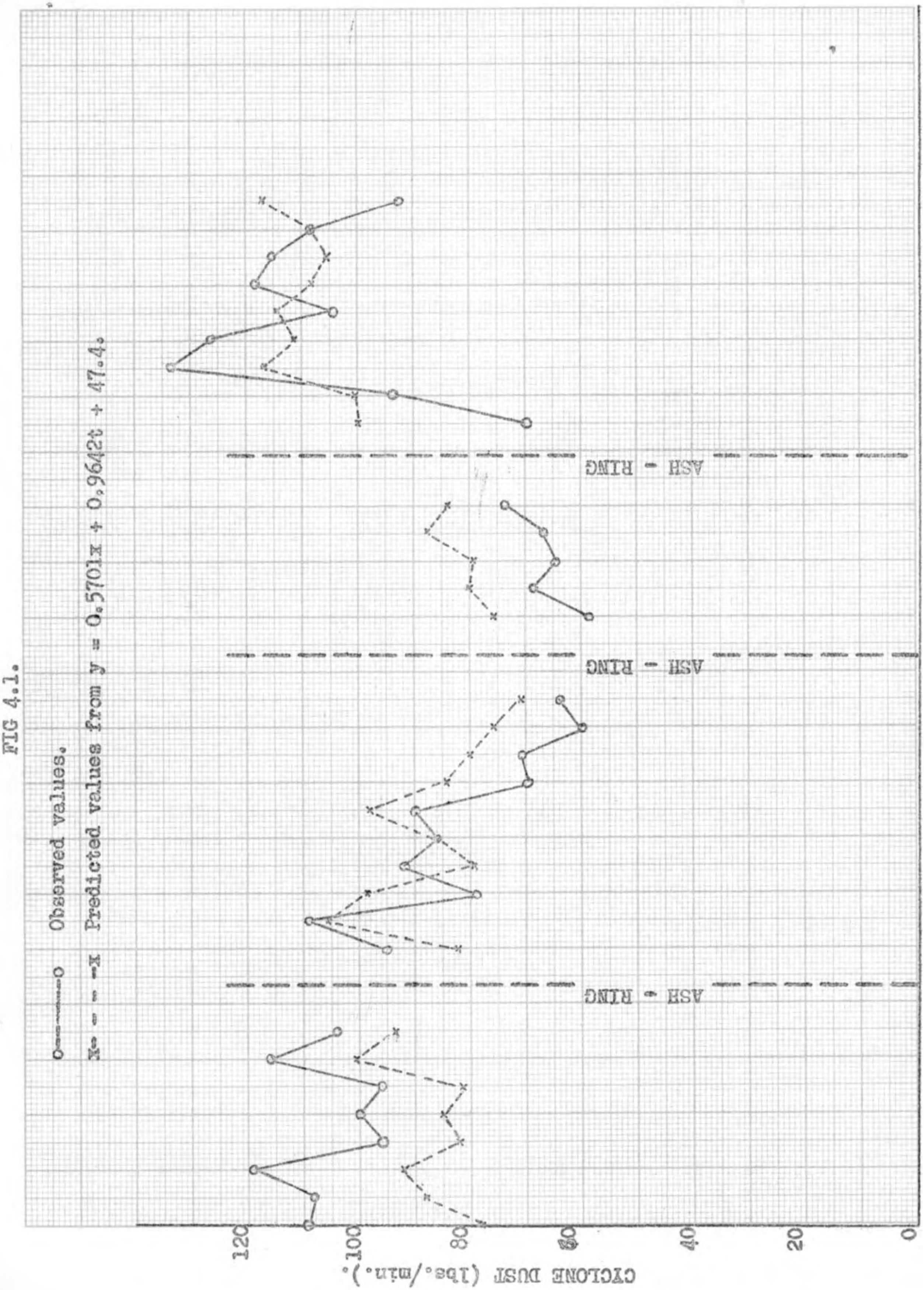


FIG 4.1.
O———O Observed values.
x- - - -x Predicted values from y = 0.5701x + 0.9642t + 47.4.
ASH - RING
ASH - RING
ASH - RING
TIME
CYCLONE DUST (lbs./min.).
0
20
40
60
80
100
120

To test whether the dependence of cyclone dust on stone fines and stone hardness was the same in all four groups it was necessary to carry out separate regressions for each group. Using the data of Table 4.A. the following sets of equations were derived.

$$\text{Group I} : \left.\begin{aligned} 818b_1 + 87b_2 &= 335 \\ 87b_1 + 31.9b_2 &= 94 \end{aligned}\right\}$$

$$\text{Group II} : \left.\begin{aligned} 2768b_1 + 154b_2 &= 1593 \\ 154b_1 + 26.4b_2 &= 133 \end{aligned}\right\}$$

$$\text{Group III} : \left.\begin{aligned} 134b_1 + 26b_2 &= 105 \\ 26b_1 + 10.8b_2 &= 6.2 \end{aligned}\right\}$$

$$\text{Group IV} : \left.\begin{aligned} 788b_1 - 99b_2 &= 241.8 \\ -99b_1 + 206b_2 &= 403 \end{aligned}\right\}$$

From which we get:-

	Group I	Group II	Group III	Group IV
b_1	0.135	0.438	1.26	0.588
b_2	2.58	2.47	- 2.46	2.24

The analyses of variance for the four groups are shown in Table 4.D.

TABLE 4.D.

ANALYSIS OF VARIANCE FOR FOUR GROUPS

Variation	Group I			Group II			Group III			Group IV		
	DF	SS	MS	DF	SS	MS	DF	SS	MS	DF	SS	MS
Regression	2	287.8		2	1025.8		2	117.2		2	1173.81	
Residual	5	234.2	46.8	7	1132.6	161.8	2	3.6	1.8	6	1817.75	303
Total	7	522.0		9	2158.4		4	120.8		8	2991.56	

We want now to combine these four analyses into one analysis but it is first necessary to test whether the four residual variances are compatible.

Accordingly Bartlett's Test was applied and the value

$$\chi^2 = 8.81 \text{ on } 3 \text{ degrees of freedom}$$

was determined. The value is just significant at the 5% level. Now Bartlett's Test is very susceptible to data which is even slightly non-normal and coupled with the fact that there are only a small number of degrees of freedom in the estimates of residual variances it was decided that this significance could possible be spurious and so the analyses were combined as shown in Table 4.E.

TABLE 4.E.

COMBINED ANALYSIS OF VARIANCE.

Variation	DF	SS	MS
4 Regressions	8	2604.61	
Residual	20	3188.15	159.4
Total	28	5792.76	

To determine the pooled estimates of the regression coefficients it was necessary to add together the four sets of equations which were used to determine the individual group regressions. This gave

$$\left.\begin{aligned} 4508b_1 + 167.8b_2 &= 2242 \\ 167.8b_1 + 275.1b_2 &= 671.8 \end{aligned}\right\}$$

from which the pooled estimates $b_1 = 0.4159$ and $b_2 = 2.1885$ were derived. The more detailed analysis of variance is shown in Table 4.F.

TABLE 4.F.

OVERALL ANALYSIS OF VARIANCE

Variation	DF	SS	MS
Regression	2	2402.83	
Differences in coefficients	6	201.78	33.63
Sum of 4 regressions	8	2604.61	
Residual	20	3188.15	159.41
Total	28	5792.76	

The variance ratio testing the differences in regression coefficients is less than unity so it was concluded that the type of dependence of cyclone dust on stone fines and stone hardness was the same in all groups, that is to say the individual regression planes for the four groups were parallel.

To test whether these planes were coincident the analyses of variance given in Table 4.B. and Table 4.E. were combined and the final analysis is as shown in Table 4.G.

TABLE 4.G.

FINAL ANALYSIS OF VARIANCE

Variation	DF	SS	MS	F
Ascribable to regression	2	5747.71	2872.86	18.03
Difference in coefficients	6	201.78	33.63	<1
Distance between planes	3	4554.08	1518.02	9.52
Residual	20	3188.15	159.41	
Total	31	13, 691.72		

The variance ratio testing the distances between the regression planes for individual groups is highly significant, greater than 0.1% level and therefore warrants further investigation.

The data should thus be represented by the four following regression planes:-

Group I : $y = 0.4159x + 2.1885t + 51.1$

Group II : $y = 0.4159x + 2.1885t + 30.3$

Group III : $y = 0.4159x + 2.1885t + 17.6$

Group IV : $y = 0.4159x + 2.1885t + 13.1$

It was not sufficient to show that the regression planes for the four groups were different, it was also necessary to try to determine why they were different. The obvious explanation was there must be some other factor, or factors, which affected the amount of dust loss from the kiln. This factor would apparently be fairly constant within each of the groups, especially in groups I and III, but would vary appreciably between the groups. Accordingly reference was made to the record sheets compiled during the test period and it appeared likely that the kiln speed might be the missing factor as it showed the sort of variation which was being looked for. Table 4.H. shows the values of kiln speed (no. of secs./ rev.) corresponding to the previous data.

TABLE 4.H.

KILN SPEED (secs./rev.)

Group I	Group II	Group III	Group IV
47.1	48.1	50.7	50.8
46.8	48.0	49.4	51.3
46.9	48.3	49.9	50.6
47.0	48.5	49.6	47.4
47.0	48.5	50.6	47.5
50.0	48.1		47.4
48.2	48.3		47.1
48.0	49.8		49.4
	50.4		48.5
	51.1		

A preliminary analysis was carried out to see whether there was a significant variation of kiln speed between the four groups as shown in Table 4.I.

TABLE 4.I.

ANALYSIS OF VARIANCE FOR KILN SPEED

Variation	DF	SS	MS
Between Groups	3	19.6558	6.5519
Within Groups	28	43.5529	1.5555
Total	31	63.2087	

The variance ratio testing between groups variation was 4.21 on 3 and 28 degrees of freedom, which is significant at the 5% level. This crude analysis showed that the variation in kiln speed between the four groups was significantly greater than the variation within the groups. Comparison of the average dust loss and the average kiln speed for each group suggested that the cyclone dust increased as the kiln speed increased i.e. as the number of seconds per revolution decreased. From a practical point of view this is quite likely as the faster the kiln rotates the more disturbance there is in the stone in the kiln and therefore a greater breakdown in the stone.

It was decided therefore to include kiln speed in the analysis and a regression plane of the type

$$(y-\bar{y}) = \beta_1 (x-\bar{x}) + \beta_2(t-\bar{t}) + \beta_3(u-\bar{u})$$

where u represents kiln speed, was set up as a model.

An overall regression ignoring any possible differences between the groups was carried out and this gave rise to the following equation:

$$y = 0.4334x + 0.8909t - 6.9581u + 394.34 \qquad \text{---------(2)}$$

An analysis of variance testing the goodness of fit of this line and the additional improvement of including the kiln speed was as follows:

TABLE 4.J.

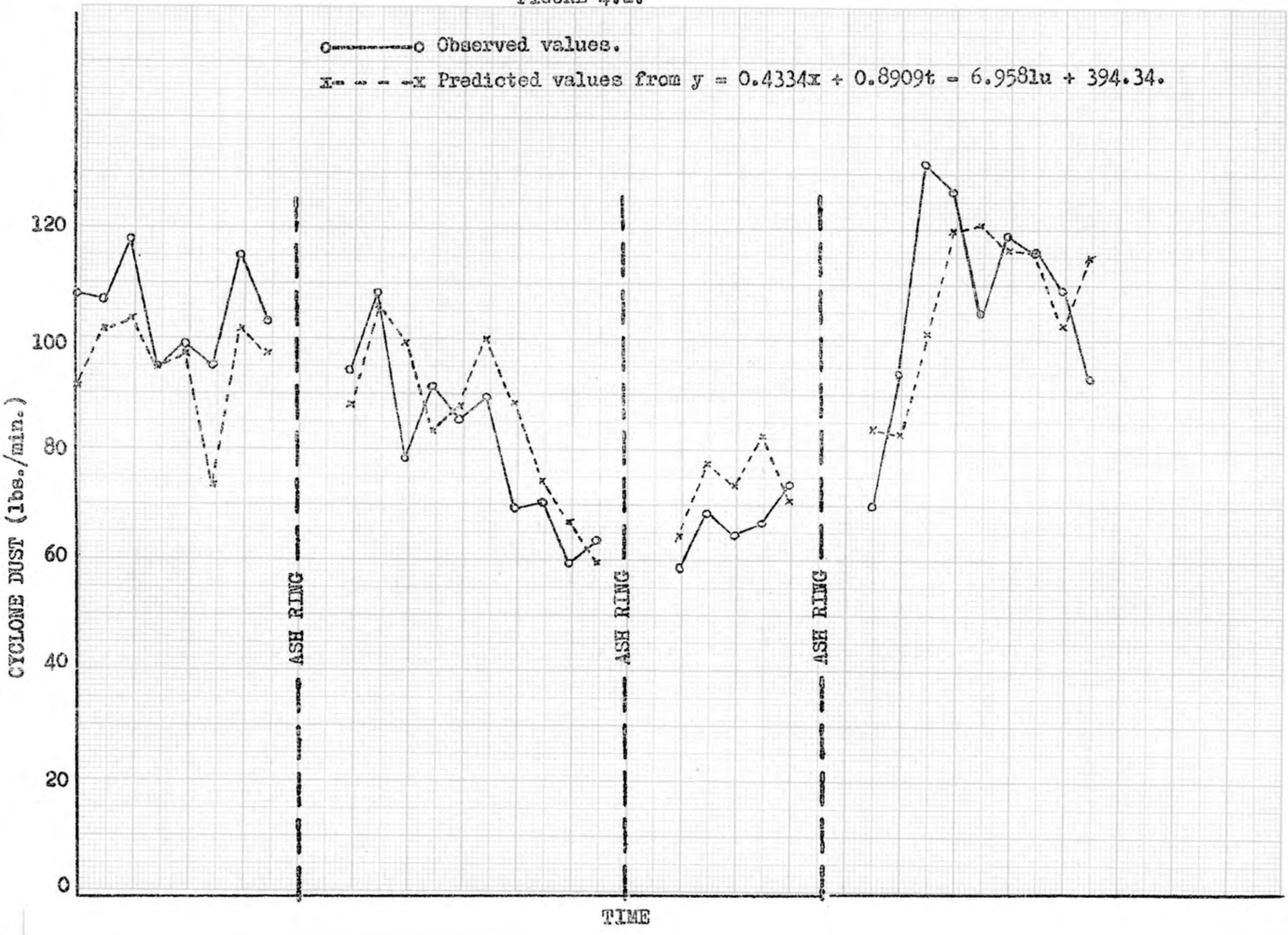
ANALYSIS OF VARIANCE TESTING LINEAR REGRESSION

Variation	DF	SS	MS
Regression (x,t)	2	5747.71	
Remainder to u	1	2897.61	2897.61
Regression (x,t,u)	3	8645.37	2881.79
Residual	28	5046.35	180.23
Total	31	13,691.72	

The variance-ratio testing the adequacy of the fit is very highly significant and so is the variance ratio testing the inclusion of kiln speed in the analysis. It appears then that the kiln speed does affect the amount of dust loss from the kiln and the fact that the b_3 is negative verifies that the dust loss increases with the kiln speed.

A graph was then plotted of observed values and values predicted from Equation (2) and this is shown in Figure 4.2. The fitted points although in better agreement with observed values than before still exhibit a discrepancy inasmuch as the marked similarity in trend and displacement in Groups I and III are still obvious. It was decided therefore to break the analysis down as before and to test whether the regression planes for

FIGURE 4.2.


o———o Observed values.
x- - - -x Predicted values from y = 0.4334x + 0.8909t - 6.9581u + 394.34.
CYCLONE DUST (lbs./min.)
0
20
40
60
80
100
120
ASH RING
ASH RING
ASH RING
TIME

the four groups were the same. This is shown in Table 4.K.

TABLE 4.K.

FINAL ANALYSIS OF VARIANCE.

Variation	DF	SS	MS
Overall regression	3	8645.37	2881.79
Differences in coefficients	9	400.96	44.55
Differences in distances	3	1841.43	613.81
Residual	16	2803.96	175.25
Total	31	13,691.72	

The variance-ratio testing the differences in the regression coefficients is less than unity while the variance-ratio testing the differences between the planes is significant at the 5% level. It was concluded therefore that although the regression planes were still parallel for the four groups they were also still displaced.

The situation was still therefore very much as it was before insofar as the differences between the groups were concerned and accordingly it seemed likely that there must be some other factor(s) affecting the dust loss from the kiln which had not yet been considered. In a further discussion on what the factor(s) might possibly be it was decided to investigate the amount of coal used during the test period. This had fortunately been recorded for one of the other tests which had been carried out at the same time.

It had been noticed on looking through the record sheets that the rate of feeding the stone into the kiln had not been constant throughout the test period, though the variation had only been small and it was decided that it would be better to change the units of measurement to allow for this. Accordingly the amount of stone fines in the feed to the kiln and the amount of dust loss from the kiln were now transformed to percentages of the total feed to the kiln. It had been assumed earlier that the amount of partly calcined material in the dust loss was negligible but chemical analysis of the dust loss samples showed that this was not true. It was possible however from this chemical analysis to convert the measured dust loss into actual dolomite and therefore as a true percentage of the feed to the kiln.

The revised observations are shown in Table 4.L.

TABLE 4.L.

REVISED OBSERVATIONS

	Stone Fines % X	Stone Hardness T	Kiln Speed (secs./rev.) U	Coal Consumption (coal worms revs. min.) V	Dust Loss % Y
GROUP I	3.2	17	47.1	113.9	17.27
	5.4	17	46.8	119.6	17.30
	6.3	17	46.9	107.0	19.09
	4.7	14	47.0	109.3	15.58
	5.4	14	47.0	113.8	16.22
	5.1	14	50.0	99.4	17.81
	7.7	19	48.2	110.5	19.02
	6.2	19	48.0	105.8	16.26
			ASH RING		
GROUP II	4.2	17	48.1	131.0	15.10
	9.2	17	48.0	117.5	17.27
	7.1	19	48.3	122.2	12.29
	3.9	15	48.5	121.0	15.43
	5.4	15	48.5	125.0	14.65
	8.0	15	48.1	117.3	15.45
	5.3	14	48.3	114.1	11.93
	4.5	14	49.8	126.9	12.55
	3.5	14	50.4	126.5	10.51
	2.4	14	51.1	116.4	11.32
			ASH RING		
GROUP III	3.2	15	50.7	132.6	9.75
	4.3	15	49.4	127.0	11.56
	4.0	16	49.9	129.6	11.27
	5.1	19	49.6	128.4	11,96
	5.1	16	50.9	134.4	13.61
			ASH RING		
GROUP IV	8.2	22	50.8	115.3	12.86
	6.7	29	51.3	107.4	17.60
	9.7	28	50.6	117.9	21.82
	6.3	38	47.4	112.0	20.80
	7.0	37	47.5	118.9	17.13
	5.9	36	47.4	121.3	19.27
	6.2	32	47.1	114.9	19.06
	6.9	32	49.4	117.7	17.93
	8.4	34	48.5	126.1	15.21

It was again suspected that if coal was going to have an effect then the variation between the four groups would be greater than the variation within the groups. An analysis of variance was carried out initially to confirm this, as shown in Table 4.M.

TABLE 4.M.

ANALYSIS OF VARIANCE FOR COAL.

Variation	DF	SS	MS
Between Groups	3	1428.89	476.30
Within Groups	28	799.92	28.57
Total	31	2228.81	

The variance-ratio testing between the groups variation is 16.67 on 3 and 28 degrees of freedom which is significant at the 1% level. Comparison of the means of coal consumption and the average percentage loss from the kiln for each of the four groups suggested that the percentage lost from the kiln decreased as the coal increased.

It was decided therefore to include coal consumption in the analysis and a regression plane of the type

$$(y - \bar{y}) = \beta_1(x - \bar{x}) + \beta_2(t - \bar{t}) + \beta_3(u - \bar{u}) + \beta_4(v - \bar{v})$$

was estimated. This gave rise to the following equation:-

$$y = 0.3740x + 0.1525t - 0.5859u - 0.150w + 56.62 \quad \text{-------} \ (3)$$

The overall analysis of variance testing the significance of this regression plane, ignoring any possible differences between the groups, is shown in Table 4.N.

TABLE 4.N.

ANALYSIS OF VARIANCE FOR OVERALL REGRESSION

Variation	DF	SS	MS
Due to Regression	4	215.1245	53.781
Residual	27	98.6001	3.652
Total	31	313.7246	

The variance-ratio testing the adequacy of the fit of the regression plane is 14.7 on 4 and 27 degrees of freedom which is very highly significant. The observed values of percentage of feed lost from the kiln and the corresponding values predicted from equation (3) are shown in Figure 4.3. The discrepancies between the observed and predicted values are no longer so obvious, visually, as they had been before but it was nevertheless decided to carry out an analysis to see whether or not the regression planes for the four groups were the same.

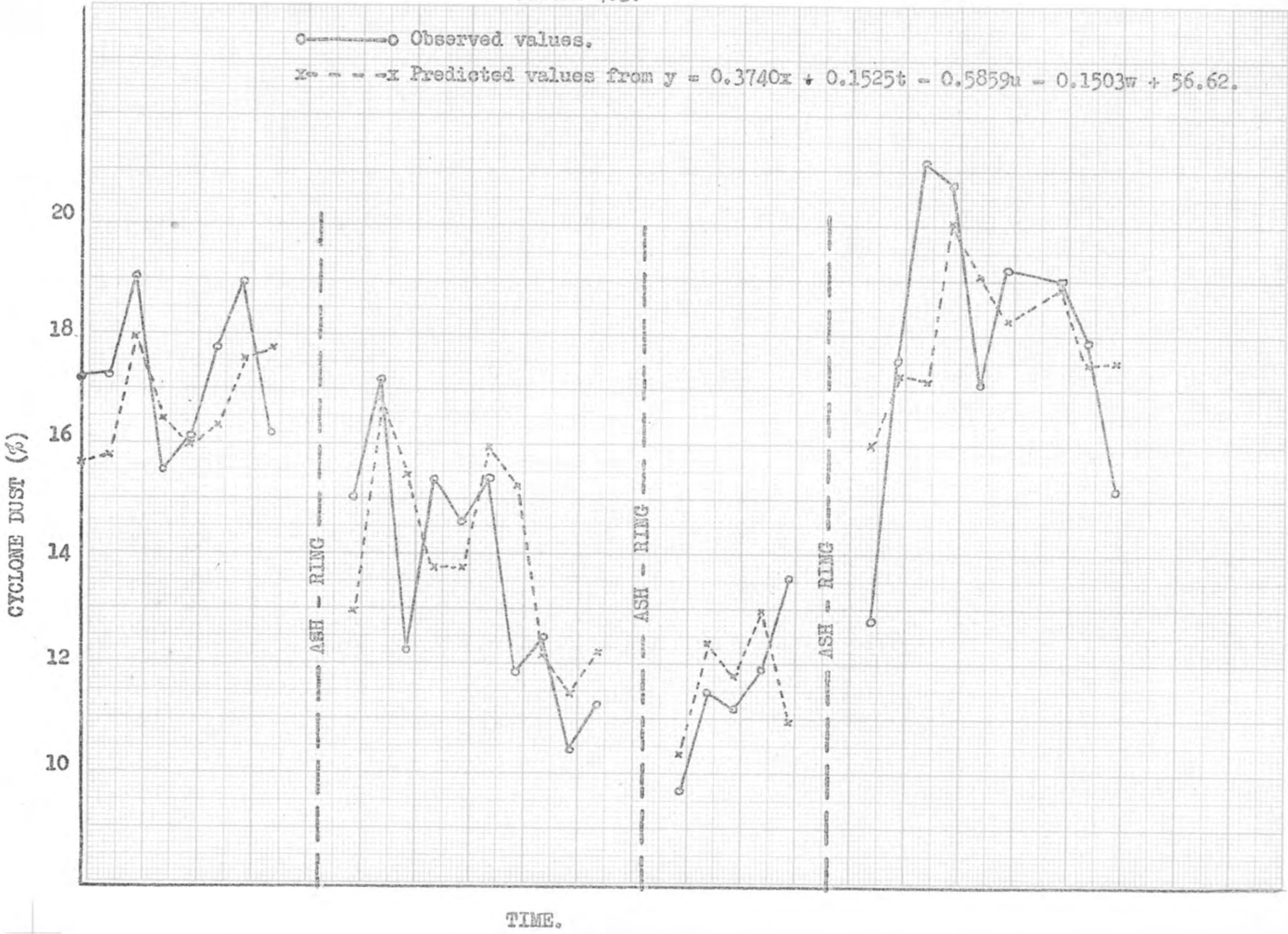
Following the same procedure as before the analysis testing group differences was carried out as shown in Table 4.O.

TABLE 4.O

FINAL ANALYSIS OF VARIANCE

Variation	DF	SS	MS
Regression	4	215.1245	53.781
Differences in Coefficients	12	28.1097	2.343
Differences in Distances	3	12.3689	4.123
Residual	12	58.1215	4.843
Total	31	313.7246	

FIGURE 4.3.


o———o Observed values.
x- - - -x Predicted values from y = 0.3740x + 0.1525t - 0.5859u - 0.1503w + 56.62.
20
18
16
14
12
10
CYCLONE DUST (%)
ASH - RING
ASH - RING
ASH - RING
TIME.

The variance-ratio testing the differences between the regression coefficients and the distances between the regression planes is less than unity in both cases. It was then concluded that the overall regression equation, ignoring any possible differences between the groups, was valid as a representation of the data.

By calculating the inverse matrix of the coefficients of the equations used to estimate the regression coefficients it is possible to determine estimates of the standard errors of the four regression coefficients. As there was no evidence that the four groups differed it was decided to use the residual variance of Table 4.N., namely 3.652 on 27 degrees of freedom as the estimate of σ^2.

This procedure gave rise to the following:

$$b_1 = 0.3740 \pm 0.2260$$

$$b_2 = 0.1525 \pm 0.0500$$

$$b_3 = -0.5859 \pm 0.2535$$

$$b_4 = -0.1503 \pm 0.0438$$

It can be seen that the standard error of b_1 is large compared with the value of b_1 but it nevertheless cannot be concluded that b_1 does not differ significantly from zero as the estimated regression coefficients are not independent of one another. The correlations between the pairs of regression coefficients were determined from the inverse matrix and it was seen that a negative correlation existed between b_1 and b_2 which was highly significant. No other correlations were significant. This highly negative correlation between b_1 and b_2 indicates that if b_1 is an over-estimate of the true value then b_2 is probably an under-estimate, and vice versa. Considering that there was very little variation in stone hardness over most of the test period and that it was only in the final stages when the level of stone hardness distinctly changed, it is feasible to assume that b_2 is over-estimated and correspondingly b_1 is under-estimated.

As stated earlier, the dust loss from the kiln was not made up solely of dolomite dust but also contained a certain amount of dolime and partly calcined material. From a chemical analysis of the dust it was possible to assess the relative amounts of dolime and dolomite plus partly calcined material, both being expressed as a percentage of the stone feed to the kiln. Statistical analyses, similar to that described above, were carried out on the dust loss from the kiln in the form of dolime and also on the dolomite plus partly calcined material, and it was seen that the latter was significantly affected by the four factors but the dolime loss was independent of the amount of stone fines fed into the kiln.

4.3. CONCLUSIONS

It has been shown that over the test period the amount of dust loss from the kiln, measured as a percentage of the stone feed to the kiln, was directly dependent upon:-

(a) Stone fines in feed, - percentage - 100 B.S. mesh (x)

(b) Stone hardness - percentage breakdown in standard milling Test (T)

(c) Kiln Speed - no. of secs./revolution (U)

(d) Coal consumption - revs./min. (V)

The relationship may be expressed mathematically as follows:-

$$y = 0.3740x + 0.1525t - 0.5859u - 0.1503v + 56.62$$

where y is the estimated percentage of the feed which is lost from the kiln.

The dust from the kiln was shown from chemical considerations to be composed of dolomite, partly calcined material and dolime. Statistical analysis showed that the amount of dust lost from the kiln as dolome depended upon stone hardness, kiln speed and coal consumption but not on stone fines while the remainder, dolomite plus partly calcined material, was significantly dependent on all the four factors.

This implies that as the amount of stone fines in the feed to the kiln increases the amount of dolomite plus partly calcined material in the dust loss from the kiln increases, whereas the amount of dolime in the dust loss is unaffected. It had been expected that an increase in stone hardness would cause a reduction in the amount of dust loss from the kiln and this was borne out in the analysis. An increase in the kiln speed caused an increase in the dust loss and this again is what one would expect as the faster the kiln rotates the more movement there is of the material in the kiln and consequently more breakdown. The affect of the stone hardness and the kiln speed was the same for both dolime and dolomite plus partly calcined material. The effect of increasing the coal consumption was to cause a marked decrease in the amount of dust loss. This is due to the fact that the stone is more quickly calcined and as calcined material is much harder than dolomite there is less breakdown in the kiln and as a result the percentage of the feed lost from the kiln decreases. Also the calcined material itself becomes substantially harder as the amount of coal used is increased and this gives a corresponding reduction in breakdown in the kiln.

The actual effect of all the factors can be assessed from a consideration of the appropriate regression coefficient and the units of measurement of each of the factors.

Having shown that these factors do affect the amount of dust loss, the question arises what can be done about it? Stone hardness cannot be controlled as the stone must be used as it comes, irrespective of whether it is hard or soft. The amount of stone fines arriving at the kiln depends on the efficiency of the initial screening and the hardness of the stone. The stone hardness is uncontrollable and the screening is in general as efficient as could be hoped for in practice so that there is little opportunity of making a saving in the dust loss from a consideration of either of these factors.

The remaining two factors, kiln speed and coal consumption, can both be varied and controlled within certain limits. In the case of coal these limits are governed by certain product quality specifications which must be maintained, i.e. there must be sufficient heat to drive off the carbon dioxide whilst at the same time there must not be too much heat or else the slakeability of the dolime would fall outside the specification. It was seen that over the test period the whole range of coal consumption, for which these physical properties were maintained, was actually covered. Bearing in mind that 1 rev./min. of the coal worms is equivalent to a coal consumption of 0.7 lb/min., it can be shown from a consideration of the regression equation (3) that 1% of the feed can be saved, other factors remaining constant, for an increase in coal consumption of 4.7 lbs/min. The average feed rate to the kiln is about 20 tons/hour so that broadly speaking 4 cwt. of dolomite could be saved by using an extra $2\frac{1}{2}$ cwt. of coal. Now the market price of dolomite is about 11 shillings per ton whereas for coal it is £5 per ton, so that an expenditure of an extra £5 for coal would result in a saving of approximately 17 shillingsworth of dolomite. It would therefore be inadvisable to put on extra coal in order to decrease the dust loss from the kiln and instead the minimum amount of coal, for which the target figures for carbon dioxide content and slakeability would still be maintained, should be used.

The indication of the analysis is that the kiln speed should be slowed down to reduce the amount of dust loss from the kiln, but here again there is a limit to how far the speed can be reduced. The minimum kiln speed is determined by the amount of material the kiln will pass and still maintain the set production figure. However it seems likely that a saving can be made here as it does not appear that the kiln speed has hitherto been governed very exactly.

CHAPTER 5.

THE ESTIMATION OF MISSING OBSERVATIONS.

In the course of carrying out experiments which are to be analysed statistically, complications often arise when some of the observations are "missing". The most common difficulty is when one or more of the experimental units are accidently lost, but the situation also occurs in which the validity of some observation(s) is suspect due to some disturbing outside factor. This latter problem arose in connection with the investigation of the rotary kiln at Hartlepool, (see Chapter 2), where the last three recordings were to be distrusted to some extent due to the large time lag which existed between them and the rest of the experiment. Consideration was given to omitting these values and deriving the appropriate least squares estimates for them in order to retain the balance of the design and thus make the computation easy. During this work a technique was developed for estimating a missing value, in two level factorial experiments, which, although not very original in principle, has the distinction of being as quick, or quicker than the already established methods. Before explaining this method it is proposed to review some of the present means of dealing with this problem of missing data.

There is also the case where an observation may be rejected because it falls outside reasonable limits and this leads to problems of truncation. This situation is not strictly relevant to the work which follows, which is essentially on methods of estimating missing data.

5.1. HISTORICAL

In the statistical analysis of any experiment a mathematical model is postulated from which it is possible to derive the least squares normal equations, which, when the experiment is complete, allow the appropriate tests of significance to be carried out very simply. When certain of the observations are missing, the orthogonality or symmetry of the design is disrupted and the correct procedure is to write down the mathematical model for all the observations that are present and then construct the least squares normal equations, which are now more difficult to solve. This is in fact the method of fitting constants as explained by Yates (20). Complications arise in the analysis of variance due to the effect of the various factors becoming entangled so that for a Randomized Block experiment, for example, the treatments sum of squares would have to be computed after allowing for block effects. This method, sometimes referred to as the "correct least squares procedure", would prove a troublesome task to many experimenters, who, although quite competent in the analysis of a complete set of data, may find difficulty when faced with an unbalanced design.

Before any methods for dealing with incomplete data were developed, the occurrence of a missing observation or a ruined plot in an experiment usually meant that the column, row or treatment containing the missing unit had to be totally omitted in order that the experiment could be analysed. This was a very wasteful course of action as much valuable information was sacrificed. The problem of extracting the maximum amount of information from the available data was first tackled by Allen and Wishart (21) for the case of a single missing plot occurring in a Randomized Block experiment or a Latin Square arrangement. Their method was to "apply the linear law which is the foundation of the analysis of variance procedure" to derive an estimate, this being the well-known technique of fitting constants. For example, in a Randomized Block experiment the linear law states that $Y = b_p + t_q$, where Y is the supposed true deviation from the mean yield of the plot having treatment q in block p. They proposed setting up the model $Y = b_p + t_q + k$ where k is a constant throughout, and put $b = 0$ and $t = 0$ for the plot which was absent. This means that k is actually the estimated value of the missing unit. The constants, in particular k, are determined by minimising $(y-b_p-t_q-k)^2$, which is of course minimising the residual sum of squares. A proviso in the final analysis of variance table is that the number of degrees of freedom for total and error is decreased by one due to fitting the value k.

Much of the basic theory for dealing with missing data has been derived by Yates (22) who, following a suggestion by R.A. Fisher, proved rigorously that a simple solution could be effected by minimising the error variance obtained when fictitious values were substituted for the missing yields. For the case of a single missing observation for which the value x is substituted, calculation of the analysis of variance leads to an expression for the error sum of squares of the form

$$Ax^2 + 2Bx + C,$$

where A,B,C are constants determined by the type of design and the values of the observations, and A is always positive. The value of x for which this expression is a minimum, is easily seen to be $x = -\ ^B/_A$. When several values are absent, substitution of the algebraic symbols x,y,z, ... lead to a set of simultaneous equations the solutions of which are the best linear estimates of the missing experimental units. For the particular cases of Randomized Block experiments and Latin Square arrangements Yates showed that simple formulae existed from which the value of x could be evaluated without going through the whole process of calculating the residual sum of squares.

The formulae given by Yates for estimating the value of a single missing unit in a Latin Square experiment is :-

$$x = \frac{r\,(R + C + T) - 2G}{(r-1)(r-2)}$$

where R,C,T are respectively the totals of the rows, column and treatment containing the missing value and G is the grand total. This formula agrees with that derived by Allen and Wishart except that an error in sign had been made in the latters' derivation.

Yates suggested that it would be convenient for the case in which several observations are absent to use repeated applications of the formula for a single missing value and to substitute approximate values for the remaining missing data. This is solving the simultaneous equations by the Gauss-Seidel iterative procedure. The analysis is then performed in the usual way, using the fictitious estimates, and the number of missing observations is then subtracted from the total and the residual degrees of freedom.

He showed that when the above technique was adopted, the treatment sum of squares is always over-estimated but may be corrected by subtracting the bias. He gives a generalized formula for the bias in a Randomized Block experiment and for the bias in a Latin Square when a single plot is absent, but in practice this bias is usually small enough to be neglected. He also gives formulae dealing with the variance of treatment means with missing units.

In two later articles Yates (23)., (24) describes methods for dealing with Latin Squares in which one or more treatments, rows or columns are missing; the principle of these methods being that of fitting by constants. In the first of these two papers he deals with the case of one missing row, column or treatment and he showed how the normal equations could be made orthogonal without any difficulty. For example, when one row is missing, row effects are orthogonal to column and treatment effects, while the normal equations for the c_i's and t_i's, where c and t represent column and treatment constants respectively, fall into pairs and can thus be solved easily. He showed that incomplete latin squares of this type give rise to unbiased estimates of error and are therefore valid experiment designs. Youden later generalised the situation to one in which several rows (or columns) are absent and thus gave rise to what are now known as Youden Squares. For the case when several treatments are missing, the normal least squares equations are such that the rows and columns are no longer orthogonal.

A method of estimating missing values, which is theoretically identical with that of Yates, was introduced by Bartlett (25) in which an analysis of covariance is carried out on a number of pseudo-variates, x_1, x_2...., each of which takes the value unity for one of the missing units and zero for the remaining units. The estimates of the missing values are then given by the appropriate regression coefficient. For example, in an experiment in which a single plot is absent let a be substituted for the missing observation as a first approximation. From the analysis of covariance table the regression coefficient, b, of y on x, is calculated from the error sum of squares and the sum of products. The value which minimises the error variance is then given by (a-b).

Anderson (26) developed a formula for missing observations in split-plot experiments by minimising the error variance. He used an analysis of covariance to derive his formula, which, also facilitated the estimation of the bias in the treatment sums of squares. Consider a split-plot experiment with α whole plot treatments, β sub-plot treatments and r replications. Suppose that the result for a single sub-unit is absent and this sub-unit receives the treatment combination $(a_i b_j)$ in the K^{th} replication. Let (A_i) and (B_j) be the total yield of all existing units with treatments, (a_i) and (b_j) respectively, and let $(A_i B_j)$ be the total of all sub-units with the treatment combination $(a_i b_j)$. Let $(R_k A_i)$ be the sum of the remaining observations in the K^{th} replicate which receive the treatment a_i.

Let the pseudo-variate, x = -1 and y = 0 for the missing value and let x = 0 and y be the actual yield for the existing units. Then the best estimate, which minimises the sub-plot error, is simply the regression coefficient of y on x, which is equal to

$$\frac{r.(R_k A_i) + \beta(A_i B_j) - A_i.}{(r-1)(\beta-1)}$$

This is of course identical in principle to Bartlett's general technique and is a special case when the value zero is substituted for the missing datum, in the initial analysis.

For the case when a whole-plot is absent, Anderson recommends that the analysis be made on a Randomized Block basis using the same missing-plot formula, and he also draws attention to the fact that the analysis of the sub-plots is unaffected in this instance.

The question of incomplete data in confounded factorial designs has been investigated by Cochrane (27). For the case of 2^m and 3^m factorial designs in which one or more effects are completely confounded, and the same design is repeated in all replications, the problem is essentially the same as for Randomized Blocks and is therefore relatively simple. When the confounding is only partial, however the situation becomes much more complex. Cochrane has developed a formula which is valid when a different arrangement of the plan is used for each replication of the experiment and when no treatment comparison is confounded more than once. In this latter case the estimate of the missing yield is given by

?

$$y = \frac{k.t(r-1)\ T + tr(r-1)\ B + KG + tV - kR - t(r-1)\ U - tr.\ Sb.}{(r-1)\left\{ t(r-1)(k-1) - (t-k) \right\}}$$

where T = total yield of all other units having the same treatments as the missing unit.

B = total yield of all other units in the same block as the missing unit.

R = sum of all existing yields in the same replication as the missing unit

Sb = sum of all the blocks which contain the same treatmentsas the missing unit in the other replications.

G = total of all existing yields in the experiment.

U = sum of the observations for all other treatments which appear in the same block as the missing treatment unit.

V = sum of the observations for all other treatments which appear in the same block as the missing treatment unit in the other replications.

t = number of treatments.

r = number of replications.

k = number of units per block.

In experiments in which the partial confounding is unbalanced and in single replications of 2^m and 3^m factorial designs, Cochrane recommends the application of the general methods of Yates (22).

Methods of dealing with the various incomplete block experiments, when some results are absent, have been developed by Cornish and these with the exception of two special cases are rather complex. These particular instances arise (a) when the incomplete blocks are ineffective so that the analysis reduces to one of the randomized blocks, or, (b) when the variation between blocks is so large that the inter-block information is negligible. In this latter case the correct procedure is to minimise the intra-block error variance.

Cornish has developed formulae for a single missing value in incomplete randomized blocks (28), and in square lattices, cubic lattices and lattice squares (29). In every case minimising the intra-block mean square was the criterion for estimating the absent datum, but, with the exception of the first-mentioned design, the formulae are necessarily complicated. In two subsequent papers (30), (31), he investigates the problem of missing sets of data, i.e. a whole block or a treatment, in balanced incomplete blocks and lattice squares. Again he minimises the intra-block variation and the method of deriving the appropriate estimates is that of fitting by constants. In his last two papers, (32), Cornish deals with lattice square experiments in which the inter-block information is also used to assess the error variance but here the situation becomes very complex indeed.

It has been shown that in general two estimates are required for the missing plot, and correspondingly two analyses of variance; one to obtain the correct inter-block error mean square, and one to get the intra-block estimate of error. In order to reduce the labour involved in using this involved procedure it has been decided to calculate the intra-block estimate only. This has been shown to serve the purpose quite well and does not bring in the bias that could possibly be introduced by the other estimate. The worst possible situation which could arise would be when the blocks are ineffective, and the problem is thus one of randomized blocks.

With the introduction of automatic computers, new ways of dealing with incomplete results had to be devised. The techniques already described are not particularly suitable to these modern machines as they involve algebra rather than arithmetic, or else they would require a different programme for every design, so that too much time and effort would be expended on a small proportion of the data unless the procedure for dealing with missing observations is standardised.

In a paper by Tocher (33), the idea was put forward of considering the difficult experimental designs in matrix notation, this system being particularly adept when the analysis is performed on an automatic computer. It also facilitates the estimation of missing data as only a slight extension of the basic programme is needed. The general method of attack is to substitute zero values for the missing observations and to estimate the corresponding expected values from an application of the basic programme. These estimates are then "corrected" by a matrix factor, which naturally depends upon the experimental design, and the experiment finally analysed with the corrected values substituted for the missing data. When only one result is absent the multiplying matrix reduces to a scalar and Tocher suggests it may be quicker when several plots are missing to apply the procedure for a single missing observation repeatedly. This technique is equivalent to deriving and solving the simultaneous equations obtained when the residual variances is minimised.

A general missing plot formula was proposed by Hartley (34) for the case of a single missing value. Three values a_0, a_1 and a_2 , unity apart, are substituted in turn for the absent result and the corresponding error sums of squares, Q_0, Q_1, Q_2 derived from the basic programme.

The best linear estimate is then given by

$$a = a_1 + (Q_0 - Q_2) / 2(Q_0 - 2Q_1 + Q_2)$$

which is the value at which the parabola through the three points (a_0, Q_0), (a_1, Q_1) and (a_2, Q_2) attains its minimum value. For several missing plots the customary iterative procedure is recommended. This method is fairly satisfactory for one absent result but for several missing units it will be inefficient as three applications of the basic programme are necessary for each missing value in every cycle of iteration.

Healy and Westmacott (35), put forward another technique for use on automatic computers. This is based on the fact that the supplied items are their own expected values derived from the least squares estimates of the block, treatment, etc., constants, so that the residuals for these units are zero. In their use of electronic computers it has been common practice for them to evaluate the residual for each plot as a comprehensive check on their calculations. When missing data occur therefore, guessed values are inserted from which the corresponding residuals are subtracted to give more accurate estimates of the absent data. The process is then repeated until the residuals for the ficticious values are zero. This technique is only a little slower than Hartley's for a single missing value, and considerably quicker when more of the data is missing.

5.2. MISSING DATA IN TWO-LEVEL FACTORIAL DESIGNS.

It can be seen from the previous section that there are several methods for estimating missing observations. For many experimental designs formulae have been derived for use when a single result is missing but, rather surprisingly, no such formulae exist for Factorial designs. The problem of missing data arose in this instance in a single replication of a 2^4 Factorial experiment. The method commonly used in such a case, see Quenouille (36), is one in which a guessed value is substituted and a better estimate derived from the basic model. This new estimate is then used as a better approximation, and the process repeated until the estimates attain steady values.

This iterative procedure can be laborious in practice and another method is described in this section, which is similar to that of Tocher (33), but which was derived independently. This method is used to obtain a formula for estimating a single missing result in a two-level Factorial experiment, and is then extended to all complete balanced block designs to give an alternative method of deriving known formulae.

SINGLE MISSING RESULT IN A TWO-LEVEL FACTORIAL EXPERIMENT.

Consider a two level Factorial experiment in which there are N trials and suppose it is required to estimate L(< N) constants, corresponding to the mean, main effects, and interactions. The (NxL) design matrix, X, will consist of L mutually orthogonal column vectors, each element of the first column being unity and the elements of the remaining columns being $\pm$ 1. If Y is the (N x 1) vector of observed responses then the least squares estimate of the vector of predicted responses is given by

$$\hat{Y} = X(X'X)^{-1}X'Y \qquad \text{--------(1)}$$

But X'X = NI, where I is the unit matrix, so that equation (1) reduces to

$$\begin{aligned} Y &= 1/N. \quad XX'Y \\ &= 1/N. \quad TY, \quad \text{say.} \qquad \text{----- (2)} \end{aligned}$$

The diagonal elements of the matrix T are the sums of squares of the elements in the rows of X and are clearly all equal to L. If $y^{*}p$ is the guessed value for the missing observation in the p^{th} trial and $\hat{y}p$ is the expected value, it follows that

$$\hat{y}p = \frac{L}{N} \quad yp + C, \qquad \text{----- (3)}$$

where C is a constant depending on the experimental design and the observed responses. The least squares estimate of the missing result, y, is its own expected value, so that we have

$$y = \frac{L}{N} y + C$$

$$\text{i.e.} \quad y = \frac{N.}{N-L} C \qquad \text{------ (4)}$$

But C is the expected response when zero is substituted for the missing result. Hence the best estimate of the missing value can be found by substituting zero for the missing datum and multiplying the expected response by N/N-L, thus dispensing with the need for successive approximations.

This result applies to the fractional factorial designs and to the multifactorial designs of Plackett and Burman, as well as the complete factorials.

Following from this, it is possible to find a second formula for estimating the missing result in two cycles of iteration. If y_1 is the expected response when y_0 is the initial guessed value for the missing observation, we have from equation (3),

$$y_1 = \frac{L}{N} y_0 + C \qquad \text{------ (5)}$$

Similarly, if y_1 is then substituted the expected value, y_2, is given by

$$y_2 = \frac{L}{N} y_1 + C \qquad \text{------ (6)}$$

Eliminating C, L and N from equations (4), (5) and (6) we have

$$y = \frac{y_0 y_2 - y_1^2}{y_0 - 2y_1 + y_2} \qquad \text{----- (7)}$$

which is the same answer as that given by applying the exponential extrapolation formula to the three points y_0, y_1, and y_2.

EXTENSION TO THE BALANCED COMPLETE BLOCK DESIGNS.

The method of estimating missing data described above can be extended to the balanced complete block designs as an alternative general method of deriving known formulae. It is shown below how a single missing result can be estimated by substituting zero and multiplying the expected response by a constant depending on the dimensions of the design.

Consider an experiment of N trials designed to estimate $L(< N)$ constants and suppose the result of one of the trials is missing. Let there be r independent groupings, corresponding to treatments and 'blocks', and let there be (n_i-1) independent constants to estimate in the i^{th} grouping. Since the r groupings are independent this implies that N/n_i is an integer and

$$1 + \sum_i (n_i-1) = L \qquad \text{------ (8)}$$

after allowing for the estimation of the mean.

Consider the $(n_i \times n_i)$ orthogonal matrix, S_i, with the elements of the first column all equal to unity such that

$$S_i S_i' = S_i' S_i = n_i . I, \qquad \text{----- (9)}$$

where I is the unit matrix. The $(n_i \times \overline{n_i-1})$ matrix, P_i, comprising the last n_i-1 columns of S_i can be used to accommodate the (n_i-1) constants in the i^{th} group. The $(N \times L)$ orthogonal design matrix, X, can then be defined as

$$X = \left[\begin{array}{c|c|c|c|c|c|c} \mathbf{1} & G_1 & G_2 & \ldots & G_i & \ldots & G_r, \end{array}\right] \qquad \text{-----} (10)$$

where $\mathbf{1}$ is an $(N \times 1)$ vector with each element unity, and G_i is an $(N \times \overline{n_i - 1})$ matrix in which each row of P_i is repeated N/n_i times.

Since $S_i' S_i = n_i . I$

$$= \left[\begin{array}{c} \mathbf{1}' \\ \hline P_i' \end{array}\right] \left[\begin{array}{c|c} \mathbf{1} & P \end{array}\right]$$

$$= \left[\begin{array}{c|c} \mathbf{1}'\mathbf{1} & 0 \\ \hline 0 & P_i'P_i \end{array}\right] \qquad \text{------} (11)$$

where $\mathbf{1}$ is an $(n_i \times 1)$ vector with all elements equal to unity, it follows that

$$P_i' \ P_i = n_i I \qquad \text{-----} (12)$$

Any column in G_i is orthogonal to any other column in G_i, and as the groupings are independent it follows that any column in G_i is orthogonal to any column in G_j. We have, therefore,

$$X'X = \left[\begin{array}{c|c|c|c} N & 0 & & 0 \\ \hline 0 & G_1'G_1 & & 0 \\ \hline & & & \\ & & & \\ & & & \\ \hline 0 & & & G_r'G_r \end{array}\right] = N.I \qquad \text{-----} (13)$$

Hence,

$$\hat{Y} = X(X'X)^{-1}X'Y$$

$$= 1/N. \; XX'Y \qquad \text{------ (14)}$$

By definition, $S_iS_i' = n_iI$, from which it follows that the sum of squares of the elements in any row in P_i, and consequently G_i, is equal to (n_i-1). The diagonal terms of XX' are thus clearly all equal to $1 + \sum_i(n -1) = L$.

If the observation in the p^{th} trial is missing and y_p, $\hat{y}_p$ are the guessed and expected values respectively, then

$$\hat{y}_p = L/N \; y_p + C \qquad \text{------ (15)}$$

where C is a constant depending on the experimental design and the observed responses. As with the two-level factorial designs, the least squares estimate of the missing result is found by substituting zero and multiplying the expected response by N/N-L.

As an example consider a split plot experiment in which there are r replications of α whole plots, each of which contains β sub-units. Let the treatment combination of the j^{th} sub-plot in the i^{th} whole plot in the k^{th} replication be missing and substitute zero for it. It is intended to derive an estimate for the missing datum such that the sub-plot error variance is minimised.

Let G = sum of all existing units

A_i = sum of all units receiving treatment a_i

B_j = sum of all units receiving treatment b_j

A_iB_j= total of all sub-units with treatment combination a_ib_j

R_k = total of remaining units in the k^{th} replicate

R_kA_i= sum of all units in k^{th} replicate which receive treatment a_i

The response, when zero is substituted for the missing experimental unit, is given by

$$y_1 = \frac{G}{\alpha\beta r} + \left\{\frac{B_j}{r\alpha} - \frac{G}{\alpha\beta r}\right\} + \left\{\frac{(A_iB_j)}{r} - \frac{A_i}{r\beta} - \frac{B_j}{r\alpha} + \frac{G}{\alpha\beta r}\right\} + \left\{\frac{R_k}{\alpha\beta} - \frac{G}{\alpha\beta r}\right\}$$

$$+ \left(\frac{A_i}{r\beta} - \frac{G}{\alpha\beta r}\right) + \left(\frac{R_kA_i}{\beta} - \frac{A_i}{r\beta} - \frac{R_k}{\alpha\beta} + \frac{G}{\alpha\beta r}\right) \qquad \text{-----} \ (16)$$

$$= \frac{r(R_kA_i) + (A_iB_j) - A_i}{r\beta} \qquad \text{--------} \ (17)$$

The correct least squares estimate, y, is given by

$$y = \frac{N}{N-L} \quad y_1, \qquad \text{------} \ (18)$$

Where $N = \alpha\beta r$ and $(N-L) = \alpha(r-1)(\beta-1)$, the number of degrees of freedom in the sub-plot error

$$y = \frac{\alpha\beta r}{\alpha(r-1)(\beta-1)} \cdot \frac{r(R_kA_i) + (A_iB_j) - A_i}{\beta r}$$

$$= \frac{r.(R_k.A_i) + (A_i.B_j) - A_i,}{(r-1)(\beta-1)} \qquad \text{------} \ (19)$$

which is the formula due to Anderson (26).

REFERENCES.

(1) BOX, G.E.P., and WILSON, K.B. "On the Experimental Attainment of Optimum Conditions" J.R. Stat. Soc., B,13:1, 1951

(2) HOERL, A.E. "A Technique for Optimizing Process Conditions". Paper presented at the Fourth A.S.M.E./I.R.D. Conference at Newark, Delaware, April 1958.

(3) PLACKETT, R.L., "Some Theorems on Least Squares". Biometrika, 36:458, 1949.

(4) BOX, G.E.P., and HUNTER, J.S. "A Confidence Region for the Solution of a Set of Simultaneous Equations with an Application to Experimental Design". Biometrika, 41 : 190, 1954.

(5) BOX, G.E.P. "The Exploration and Exploitation of Response Surfaces: Some General Considerations and Examples". Biometrics, 10:16, 1954.

(6) BOX, G.E.P., and YOULE, P.V. "The Exploration and Exploitation of Response Surfaces: An Example of the Link between the Fitted Surface and the Basic Mechanism of the System". Biometrics, 11: 287, 1955.

(7) READ, D.R. "The Design of Chemical Experiments". Biometrics, 10: 1. 1954.

(8) DAVIES, O.L. (Editor). "The design and Analysis of Industrial Experiments". Oliver and Boyd (Edinburgh 1953).

(9) BOX, G.E.P., and COUTIE, A.G. "Application of Digital Computers in the Exploration of Functional Relationships". Proc. Inst. Elec. Engineers. 103: 100, 1956.

(10) BOX, G.E.P. "Multi-Factor Designs of First Order". Biometrika, 39: 49, 1952.

(11) PLACKETT, R.L. and BURMAN, J.P. "The Design of Multifactorial Experiments". Biometrika, 33: 305, 1946.

(12) DANIEL, C. "On Varying One Factor at a Time". Biometrics, 14: 430, 1958.

(13) DE BAUN, R.M. "Block Effects in the determination of Optimum Conditions". Biometrics, 12: 20, 1956.

(14) HARTLEY, H.O. "Smallest Composite Designs for Quadratic Response Surfaces". Biometrics, 15: 624, 1959.

(15) BOX, G.E.P. and HUNTER, J.S. "Multi-Factor Experimental Designs for Exploring Response Surfaces". Ann. Math. Stat., 28: 195, 1957.

(16) BOSE, R.C. and DRAPER, N.R. "Second Order Rotatable Designs in Three Dimensions". Ann. Math. Stat., 30: 1097, 1959.

(17) GARDINER, D.A., GRANDAGE, A.H.E. and HADER, R.J. "Third Order Rotatable Design for Exploring Response Surfaces". Ann. Math. Stat., 30: 1082, 1959.

(18) BOX, G.E.P., and DRAPER, N.R. "A Basis for the Selection of a Response Surface Design". Amer. Stat. Assoc., 54: 622, 1959.

(19) JOWETT, G.H., "The Comparison of Means of Industrial Time Series". App. Stat., 4: 32, 1955.

(20) YATES, F. "The Principles of Orthogonality and Confounding in Replicated Experiments". Jour. Agr. Sci., 23: 108, 1933.

(21) ALLEN, F.E. and WISHART, J. "A Method of Estimating the Yield of a Missing Plot in Field Experimental Work". Jour. Agr. Sci., 20: 399, 1930.

(22) YATES, F. "The Analysis of Replicated Experiments When the Field Results are Incomplete". Emp. Jour. Exp. Agr., 1: 129, 1933.

(23) YATES, F. "Incomplete Latin Squares". Jour. Agr. Sci., 26: 301, 1936.

(24) YATES, F. and HALE, R.W. "The Analysis of Latin Squares when Two or More Rows, Columns or Treatments are Missing". Supp. J.R. Stat. Soc., 6: 67, 1939.

(25) BARTLETT, M.S. "Some Examples of Statistical Methods of Research in Agriculture and Applied Biology". Supp. J.R. Stat.Soc., 4: 137, 1939.

(26) ANDERSON, R.L. "Missing Plot Techniques". Biometrics, 2: 41, 1946.

(27) COCHRAN, W.G. and GERTRUDE COX. "Experimental Designs". Wiley (New York) 1950.

(28) CORNISH, E.A. "The Estimation of Missing Values in Incomplete Randomized Block Experiments". Ann. Eugen., 10: 112, 1940.

(29) CORNISH, E.A. "The Estimation of Missing Values in Quasi-Factorial Designs". Ann. Eugen. 10: 137, 1940.

(30) CORNISH, E.A. "The Analysis of Quasi-Factorial Designs with Incomplete Data". 1. Incomplete Randomized Blocks". Jour. Aust. Inst. Agr. Sci., 6: 31, 1940.

(31) CORNISH, E.A. "The Analysis of Quasi-Factorial Designs with Incomplete Data. 2. Lattice Squares". Jour. Aust. Inst. Agr. Sci., 7: 19, 1941.

(32) CORNISH, E.A. "The Recovery of Inter-block Information in Quasi-Factorial Designs with Incomplete Data. 1. Square, Triple and Cubic Lattices, (Bul. 158, 1943). 2. Lattice Squares, (Bul.175, 1944)". Coun. Sci. Ind. Res. (Aust).

(33) TOCHER, K.D. "The Design and Analysis of Block Experiments". J.R. Stat. Soc., B, 14: 45, 1952.

(34) HARTLEY, H.O. "Programming Analysis of Variance for General Purpose Computers". Biometrics, 12: 1956.

(35) HEALY, M.J. and WESTMACOTT, M. "Missing Values in Experiments Analysed on Automatic Computers". App. Stat. 5: 203, 1956.

(36) QUENOUILLE, M.H. "The Design and Analysis of Experiments". Griffin, (London, 1953).